Série Relatórios Metodológicos

VOLUME 25

Metodologia do Censo Demográfico 2000

IBGE
Instituto Brasileiro de Geografia e Estatística

Presidente da República
Luiz Inácio Lula da Silva

Ministro do Planejamento, Orçamento e Gestão
Guido Mantega

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE

Presidente
Eduardo Pereira Nunes

Diretor Executivo
José Sant`Anna Bevilaqua

ÓRGÃOS ESPECÍFICOS SINGULARES

Diretoria de Pesquisas
Maria Martha Malard Mayer

Diretoria de Geociências
Guido Gelli

Diretoria de Informática
Luiz Fernando Pinto Mariano (em exercício)

Centro de Documentação e Disseminação de Informações
David Wu Tai

Escola Nacional de Ciências Estatísticas
Pedro Luis do Nascimento Silva

Ministério do Planejamento, Orçamento e Gestão
Instituto Brasileiro de Geografia e Estatística - IBGE

Série Relatórios Metodológicos
volume 25

Metodologia do Censo Demográfico 2000

Rio de Janeiro
2003

Instituto Brasileiro de Geografia e Estatística - IBGE
Av. Franklin Roosevelt, 166 - Centro - 20021-120 - Rio de Janeiro, RJ - Brasil

ISSN 0101-2843 **Série Relatórios Metodológicos**
Divulga as metodologias empregadas nas diversas fases do planejamento e execução das pesquisas do IBGE.

ISBN 85-240-3701-6 (CD-ROM)

ISBN 85-240-3700-8 (meio impresso)

© IBGE, 2003

Elaboração do arquivo PDF
Roberto Cavararo

Produção da multimídia
Marisa Sigolo Mendonça

Capa
Mauro Emílio Araújo - Coordenação de *Marketing*/Centro de Documentação e Disseminação de Informações-CDDI

Sumário

Apresentação

1 Introdução

1.1 Histórico

1.2 Legislação

1.3 Periodicidade e data de referência

2 Planejamento

2.1 Avaliação das operações censitárias de 1996 e recomendações para o Censo 2000

2.2 Participantes externos

2.3 Integração com os países do Mercosul, Chile e Bolívia

2.4 Custos e orçamentos

2.5 Cronograma geral de atividades

3 Logística da operação

3.1 Organização geral

3.2 Escolha da tecnologia de captura de dados

3.3 Sistemas de controle e apuração

3.4 Centros de Captura de Dados

3.5 Impressão, distribuição e retorno do material

4 Mobilização

4.1 Concurso da logomarca

4.2 Projeto Vamos Contar!

4.3 Comissões Censitárias Municipais

4.4 Promoção local e campanha publicitária nacional

4.5 Atendimento telefônico

4.6 Censo 2000 na Internet

4.7 Revista Vou te Contar

5 Provas Piloto e Censo Experimental

5.1 Primeira Prova Piloto

5.2 Segunda Prova Piloto

5.3 Prova Piloto conjunta dos países do Mercosul ampliado

5.4 Censo Experimental

6. Desenho conceitual

6.1 Âmbito da pesquisa

6.2 Conteúdo da investigação

6.3 Conceitos

6.4 Classificações

7 Base Operacional Geográfica

7.1 Diretrizes

7.2 Elaboração da Base Operacional Geográfica

7.3 Modernização da componente gráfica

8 Treinamento

8.1 Estratégia

8.2 Metodologia e execução

8.3 Recursos didáticos

8.4 Acompanhamento e avaliação

9 Coleta das informações

9.1 Estratégias para a coleta

9.2 Amostragem na coleta

9.3 Instrumentos de coleta

9.4 Procedimentos de atuação no campo

9.5 Acompanhamento da coleta

9.6 Evolução da coleta

10 Pesquisa de Avaliação da cobertura da coleta

10.1 Planejamento da pesquisa

10.2 Amostra da pesquisa

10.3 Instrumentos de coleta

10.4 Treinamento das equipes

10.5 Coleta dos dados

10.6 Batimento e reconciliação dos dados

10.7 Métodos de estimação

11 Captura dos dados

11.1 Dados dos questionários

11.2 Dados da Folha de Coleta

12 Apuração dos dados

12.1 Resultados preliminares

12.2 Sinopse preliminar

12.3 Resultados do conjunto universo

12.4 Resultados preliminares da amostra

12.5 Resultados da amostra

12.6 Tabulação dos dados

13 Disseminação

13.1 O Censo e a mídia

13.2 Produtos e serviços

14 Base de dados

14.1 Banco Multidimensional de Estatísticas

14.2 Sistema IBGE de Recuperação Automática

14.3 Escolha do SAS

Apresentação

Através da "Série Relatórios Metodológicos", o IBGE documenta e divulga as metodologias empregadas nas diversas fases do planejamento e execução de suas pesquisas. Por meio da explicitação dos procedimentos usados e do amplo debate técnico daí proveniente, o IBGE tem como objetivo o aprimoramento dos conceitos e métodos utilizados e, portanto, de pesquisas futuras.

Dando prosseguimento a esta série, o IBGE publica neste volume a Metodologia do Censo Demográfico 2000. Os resultados do Censo 2000 revelam o retrato do país no fim do milênio e podem ser considerados como a principal fonte de informações para a análise sobre a realidade nacional, servindo para fundamentar o exercício da cidadania, as políticas públicas e os investimentos privados, entre outros.

Este volume contém uma descrição detalhada de todos os processos que envolveram a realização do Censo Demográfico 2000, desde as etapas de planejamento e organização da operação, passando pelas novas tecnologias e sistemas desenvolvidos especificamente para a pesquisa, chegando até às diversas formas de disseminação e divulgação dos resultados.

Com esta publicação, o IBGE encerra o plano editorial de divulgação do Censo Demográfico 2000, oferecendo à sociedade não somente um conjunto de resultados através de volumes impressos e arquivos em meio magnético, mas também, uma completa visão dos procedimentos técnicos e operacionais que levaram à construção de sua base de dados.

Esta publicação é composta por um volume impresso, acompanhada por um CD-ROM com um uma grande quantidade de anexos que complementam e detalham as descrições metodológicas e operacionais.

O Instituto Brasileiro de Geografia e Estatística cumpre assim o seu compromisso de Retratar o Brasil com informações necessárias ao conhecimento de sua realidade e ao exercício da cidadania.

Eduardo Pereira Nunes
Presidente do IBGE

1 Introdução

A palavra censo vem do latim census e quer dizer “conjunto dos dados estatísticos dos habitantes de uma cidade, província, estado, nação”. O Censo Demográfico, que acontece no Brasil de dez em dez anos, realizado pelo IBGE, tem por objetivo contar os habitantes do território nacional, identificar suas características e revelar como vivem os brasileiros, pois conhecer em detalhe como é e como vive a população é de extrema importância para o governo e para a sociedade.

Desde 1872, vêm sendo realizados censos no País e, de certa forma, a sua história acaba por registrar também um pouco da história do Brasil. Este primeiro capítulo da Metodologia do Censo Demográfico 2000 apresenta, inicialmente, o histórico dos censos no Brasil, permitindo acompanhar a evolução da investigação das características da população ao longo do tempo. Vale destacar a parte do texto que descreve as principais características dos censos até 1970, que foi reproduzida a partir do histórico apresentado na publicação Metodologia do censo demográfico de 1980 (1983).

Também são apresentados neste capítulo, outros dois tópicos necessários ao acompanhamento e ao entendimento dos textos sobre as etapas de realização do Censo 2000. O primeiro deles é a legislação, relacionando leis, decretos e outros documentos que serviram de base jurídica para a realização da pesquisa; o outro trata da periodicidade e data de referência da investigação das informações.

1.1 Histórico

Não obstante ser o Brasil um País com modesta tradição estatística, é certo que, desde os tempos coloniais, realiza censos de população, hoje, mais do que no passado, com objetivos mui-

to bem definidos. Segundo informa Joaquim Norberto de Souza e Silva, em *Investigações sobre os recenseamentos da população geral do Império e de cada província de per si, tentadas desde os tempos coloniais até hoje*[1] , antes de 1870 nossas estatísticas de população eram meramente conjecturais não passando de estimativas destinadas, basicamente, a atender aos interesses da Metrópole – maiores do que geralmente se supõe – preocupada sempre em conhecer os seus domínios na América do Sul.

Assim é que a recomendação às autoridades eclesiásticas, para que procedessem ao arrolamento das populações das áreas subordinadas às respectivas jurisdições, permitiu ao Abade Corrêa da Serra estimar em 1 900 000 "almas" a população do Brasil no ano de 1776. De outro lado, a "carta régia de 8 de julho de 1800, dirigida ao Vice-Rei do Estado do Brasil, enumera, entre as primeiras obrigações inerentes às funções daquele representante da coroa portuguesa em terras brasileiras, a remessa para o reino de elementos estatísticos".

Pelos dados históricos, constantes dos arquivos oficiais e particulares, o primeiro recenseamento da população do Brasil teria sido efetuado em 1808, em decorrência do aviso de 16 de março daquele ano, baixado pelo Ministro dos Negócios da Guerra, D. Rodrigo de Souza Coutinho, mais tarde Conde de Linhares. Realizado o arrolamento, foi apurada uma população total, para o País, de 4 000 000 de habitantes.

Esse arrolamento, cujos pormenores operacionais não se tornaram suficientemente conhecidos, foi realizado para atender especificamente a interesses militares, de recrutamento para as forças armadas, o que pode dar lugar a suspeitas de que seus resultados tenham ficado aquém da realidade, seja em razão da natural prevenção do povo contra as operações censitárias, seja, principalmente, em razão de seus objetivos.

Convém assinalar, porém, que a opinião abalizada de Humboldt estimou em 4 000 000 de habitantes a população do Brasil em 1810, isto é, dois anos depois, e que uma pesquisa realizada pelo Conselheiro Antônio Rodrigues Velloso de Oliveira, apurando o total de 4 396 132 habitantes para o País, em 1818, constituem elementos favoráveis de confronto ou, pelo menos, indicam não ter havido grandes deficiências no levantamento realizado sob os auspícios do Conde de Linhares, em 1808.

No juízo de Joaquim Norberto, o inquérito realizado pelo Conselheiro Antônio Rodrigues de Oliveira, em 1819, foi, todavia, o primeiro censo de população realizado no País, pela sua aproximação da verdade, pelo possível esmero de sua organização e pela fé que nele se pode depositar. Esse inquérito foi realizado após a elevação da colônia à categoria de reino, quando, pela resolução de 24 de junho de 1818, foi determinado o reestudo da divisão dos antigos bispados e a criação dos que mais necessários parecessem[2].

Convém ponderar, entretanto, que os elementos de que se serviu o Conselheiro Velloso de Oliveira, para a sua estimativa, procediam de várias fontes, sem a precisa homogeneidade, tais como informações prestadas pelos párocos, mapas enviados pelos ouvidores ao Desembargo do Paço e por outros magis-

---

[1] Memória anexa ao *Relatório apresentado ao Ministro e Secretário de Estado dos Negócios do Império...*, em 1870, pelo titular da mesma pasta, Conselheiro Paulino José Soares de Souza.

[2] *Memória estatística do Império do Brasil*, publicado na Revista Trimestral do Instituto Histórico e Geográfico Brasileiro, t. 58, pt. 1, n. 91, p. 91-99, 1895. Texto de autor desconhecido, provavelmente escrito em 1829, e oferecido ao Marquês de Caravellas.

trados à Intendência Geral de Polícia, esclarecimentos referentes a diversos anos, e abrangendo o período de 1815 a 1818. Para compensar as deficiências de levantamento, quanto aos menores de sete anos, à tropa paga e a outros grupos de população, provavelmente não recenseados, foram feitos acréscimos, no total obtido, de até 25% dos dados apurados. Desta forma, 735 607 habitantes foram arbitrariamente calculados e integrados ao total geral, juntamente à cifra de 800 000 índios, resultante de uma simples avaliação sem qualquer fundamento científico.

Proclamada a Independência Nacional, em 1822, não houve, de imediato, uma sensível mudança no que diz respeito à estatística praticada no País, notadamente quanto à que servia de base aos estudos da população. Como no tempo da Metrópole, de vez em quando um dispositivo legal ou regulamentar demonstrava as boas disposições dos governos no sentido de verificar o progresso experimentado pela população brasileira, do ponto de vista de seu crescimento e estrutura.

A Comissão de Estatística Geográfica, Natural, Política e Civil, criada por Decreto de 25 de novembro de 1829, para tratar do assunto, teve vida efêmera e nada produziu de eficaz durante o curto tempo em que funcionou[3]. Nem mesmo o célebre Regulamento nº 120, de 31 de janeiro de 1842, que no Parágrafo 17, do Artigo 18, incumbiu os chefes de polícia de fazer o arrolamento das pessoas domiciliadas nas províncias sob sua jurisdição, ou a Lei nº 387, de 19 de agosto de 1846, que determinou, no seu Artigo 107, se procedesse à realização de censos de população a cada oito anos, ou, ainda, a Lei nº 586, de 6 de setembro de 1850, que autorizou o governo a despender o que fosse necessário para levar a efeito, no menor prazo possível, o Recenseamento Geral do Império, com especificação do que respeitasse a cada uma das províncias[4], alteraram aquela situação, embora tivesse este último instrumento desempenhado um importante papel na caracterização da necessidade imperiosa de o País realizar censos demográficos de periodicidade constante.

Em circular de 23 de outubro de 1854, o Ministro dos Negócios do Império ordenou que fosse realizado um levantamento em cada província. Os elementos coligidos nesse inquérito foram divulgados em relatório, em 1856, e apontavam o total de 7 677 800 indivíduos.

Por aviso de 28 de novembro de 1867, também do Ministro do Império, os Presidentes das Províncias foram novamente instados a fornecer as informações que pudessem colher sobre as populações dos territórios sob sua jurisdição, sem que a medida intentada surtisse, ainda dessa vez, o desejado efeito. A essa época, um trabalho elaborado para distribuição na Exposição Universal de Paris registrava que a população do Brasil estaria próxima de 11 780 000 habitantes, assim distribuídos: 1 400 000 escravos; 9 880 000 habitantes livres; e 500 000 indígenas.

O *Atlas do Império do Brasil*, entretanto, divulgado em 1868, registrava um total de 10 030 000 para o conjunto do País, estimado por Cândido Mendes de Almeida. De forma semelhante, a estimativa oficial, feita em 1869 por Thomaz Pompeu de Souza Brazil, atribuiu ao Império 10 415 000 habitantes.

---

[3] O insucesso dessa Comissão foi atribuído ao defeito de sua constituição, pois que não era composta de profissionais, e os seus titulares, distraídos com outras preocupações, não consagravam a atividade precisa ao êxito dos trabalhos confiados ao referido conselho.

[4] Os Decretos nº 797 e 798, ambos de 18 de junho de 1851, mandaram elaborar os regulamentos para a organização do censo geral e do registro civil dos nascimentos e óbitos, sendo que de suas diretrizes decorreram as criações da Diretoria Geral do Censo e das Comissões Provinciais, para a execução da operação censitária de que tratava essa Lei. A operação, em vias de realizar-se, ou em pleno início de execução, foi sustada pelo Decreto nº 907, de 29 de janeiro de 1852.

Assim, nova investigação foi solicitada pelo Ministro do Império em aviso datado de 25 de janeiro de 1870. De posse de resultados aproveitáveis para 13 províncias, o Ministério encarregou Joaquim Norberto de Souza e Silva de elaborar relatório circunstanciado sobre a população do Brasil. Nesse relatório, o autor resumiu os resultados de estimativas e inquéritos feitos em diversas épocas, desde os tempos coloniais, para o conjunto do País, na seguinte tabela:

Tabela 1.1 – População brasileira - 1776/1869

Anos	Autoridades	População
1776	Abade Corrêa da Serra	1 900 000
1808	D. Rodrigo de Souza Coutinho	4 000 000
1810	Alexandre Humboldt	4 000 000
1815	Conselheiro Velloso de Oliveira	2 860 525
1817	Henry Hill	3 300 000
1819	Conselheiro Velloso de Oliveira	4 395 132
1825	Casado Giraldes	5 000 000
1827	Rugendas	3 758 000
1829	Adriano Balbi	2 617 900
1830	Malte – Brun	5 340 000
1834	Senador José Saturnino	3 800 000
1850	Senador Cândido Baptista de Oliveira	8 000 000
1856	Barão do Bom Retiro	7 677 800
1867	"O Império na Exposição, etc."	11 780 000
1868	Cândido Mendes	11 030 000
1869	Senador T. Pompeu de Souza Brazil	10 415 000

Fonte: Silva, J. N. de S. e. Investigações sobre os recenseamentos da população geral do império e de cada província de per si tentados desde os tempos coloniaes até hoje. Rio de Janeiro: Typ. Perseverança, 1870. f. 167.

Talvez o maior mérito desse relatório tenha sido o de despertar, nas autoridades, a consciência da necessidade de realização do recenseamento, em moldes amplos, garantindo seu êxito através de medidas administrativas capazes de assegurar o registro de informes fidedignos.

A Assembléia Legislativa, então, na forma prevista no Decreto nº 1.829, de 9 de setembro de 1870, criava a Diretoria Geral de Estatística por outro Decreto, nº 4.676, de 14 de janeiro de 1871. Para um fiel registro histórico, é de dever registrar que, antes mesmo da criação da Diretoria Geral de Estatística, ainda no ano de 1870, realizou-se com êxito o Censo do Município Neutro, a que se seguiu, em 1872, idêntica operação, ampliada ao conjunto do Império, de acordo com o disposto no Decreto nº 4.856, de 30 de dezembro de 1871.

O Censo Geral do Império, a que aludia o Decreto nº 4.856, foi realizado na data fixada, 1ª de agosto de 1872, e correu tão regularmente quanto possível, em quase todo o Território Nacional, com exceção das províncias de Minas, São Paulo e Mato Grosso, onde motivos imprevistos e de força maior obstaram que os trabalhos fossem executados no tempo estabelecido[5].

[5] No Relatório da Diretoria Geral de Estatística, apresentado pelo Dr. José Maria do Couto ao Ministro e Secretário dos Negócios do Império, Conselheiro João Alfredo Corrêa de Oliveira, encontram-se interessantes informações sobre a organização interna e a marcha dos serviços de apuração do Recenseamento de 1872.

O Relatório da Diretoria Geral de Estatística sobre a organização interna e a marcha dos serviços de apuração desse censo "dá uma expressiva idéia da morosidade dos processos de apuração então adotados e a cargo do reduzido pessoal responsável pelo serviço", bem como assinala "a falta do devido tirocínio na prática de um trabalho inteiramente novo".

Apesar da pobreza dos meios disponíveis, 10 112 061 habitantes foram recenseados em 1872[6], em todas as províncias, e a sua distribuição se fez segundo a cor, o sexo, o estado de livres ou escravos, o estado civil, a nacionalidade, a ocupação e a religião.

O relativo êxito do Recenseamento de 1872 devia incentivar o governo imperial a prosseguir nos seus esforços para melhorar a estatística do País, habilitando o Departamento criado em 1871 a proceder, em melhores bases, às pesquisas que lhe cumpria efetuar no campo da Demografia, objetivo que só poderia atingir com a normalização dos inquéritos censitários, levados a efeito periodicamente e em moldes cada vez mais aperfeiçoados, graças à experiência adquirida na série de operações sucessivas que tivesse a seu cargo realizar. Não perseveraram, todavia, nessa esclarecida política administrativa os estadistas do passado regime. A mudança de orientação do governo imperial devia, em breve, concretizar-se no desaparecimento da Diretoria Geral de Estatística, por força do Decreto nº 2.940, de 31 de outubro de 1879. O Recenseamento de 1880 teve sua execução transferida para 1887 e, ainda assim, ficou apenas em projeto, e os últimos anos de monarquia assinalaram a indiferença dos poderes públicos, relativamente aos serviços da extinta repartição[7].

A Proclamação da República, em 15 de novembro de 1889, veio oportunamente pôr termo a essa injustificável indiferença dos poderes públicos em matéria de estatística. Pelo Decreto nº 113-d, de 2 de janeiro de 1890, foi restaurada a extinta Diretoria Geral de Estatística. O Decreto nº 331, de 12 de abril do mesmo ano, que derrogou aquele, deu-lhe novo regulamento e melhor definiu sua competência: "dirigir os trabalhos do recenseamento geral da população, segundo programa e os mapas que houvesse organizado, e dar-lhes publicidade". O Artigo 15 do mesmo Decreto fixava o dia 31 de dezembro de 1890 para a execução do I Censo da República.

Os trabalhos relativos a esse Censo foram regulados pelas instruções mandadas executar pelo Decreto nº 659, de 12 de agosto de 1890, referendado pelo Ministro Dr. José Cesário de Faria Alvim.

O Recenseamento foi realizado utilizando-se boletins, ou listas de famílias, nos quais se perquiriam as condições da população: quanto às características dos habitantes isoladamente considerados – nome, sexo, idade, estado civil, cor, filiação, naturalidade, nacionalidade dos pais, defeitos físicos e residência; quanto à família – nome de cada um dos cônjuges, ano de casamento, número de matrimônios, grau de parentesco do casal e número de filhos vivos e já falecidos, segundo o sexo e também relativamente aos defeitos físicos; quanto à sociedade – nome, nacionalidade, relação com o chefe da casa, instrução, culto professado, profissão, título científico, literário ou artístico, e renda.

---

[6] Para efeito de registro histórico, em razão de sua maior complexidade e, sobretudo, do controle a que foi submetida toda a operação, o Censo de 1872 é considerado, através dos anos, o primeiro Censo Demográfico realizado no País.

[7] As antigas atribuições da extinta Diretoria Geral de Estatística passaram a ser exercidas por uma simples seção da Secretaria do Ministério do Império, sem qualquer poder decisório.

O inquérito foi dirigido pela Diretoria de Estatística, auxiliada por Comissões Censitárias paroquiais e agentes recenseadores, em número variável, fixado pelo Ministro do Interior, na Capital Federal, e pelos presidentes ou governadores, nos estados, de acordo com as comissões censitárias e as repartições locais de estatística, as quais tinham atribuições consultivas, podendo, contudo, avocar a direção do serviço onde essa medida conviesse à boa marcha dos trabalhos a serem realizados.

O Artigo 12, das instruções baixadas com o Decreto nº 659, definia os encargos das Comissões Censitárias e Repartições de Estatística quanto aos trabalhos que lhes cumpria superintender. Cabia a esses órgãos proceder à divisão do território das paróquias em tantas seções quantas fossem indispensáveis, para que o recenseamento pudesse ser executado escrupulosa e facilmente por um só agente recenseador; nomear esses agentes, escolhendo-os entre pessoas que soubessem ler e escrever corretamente, além de outros predicados capazes de assegurar o bom desempenho do cargo; distribuir o material censitário pelos agentes recenseadores, levando em conta, nessa distribuição, o número provável de domicílios a serem recenseados em cada seção, a fim de que as sobras verificadas em uma delas não concorressem para a deficiência de formulários em outras, daí resultando prejuízos ao inquérito.

As Comissões Censitárias locais e as Repartições de Estatística tinham ainda como atribuições a fiscalização dos serviços dos recenseadores e sua orientação na solução dos casos imprevistos e duvidosos; o exame das informações registradas nos boletins coletados, preenchendo as possíveis lacunas retificando os registros inexatos e promovendo a remessa dos questionários, expurgados dos erros e omissões, à Diretoria Geral de Estatística, remessa obrigatoriamente acompanhada de um minucioso relatório sobre os principais fatos ocorridos durante o recenseamento. Acrescente-se a tais incumbências uma lista com os nomes das pessoas que mais tivessem colaborado para o êxito do inquérito censitário.

Os trabalhos de apuração do Censo de 1890 ocorreram com grande morosidade, seja porque naquela época não eram ainda de aplicação corrente equipamentos mecânicos nesses trabalhos, seja porque concorreram para isso os acontecimentos políticos observados nos primeiros anos do regime republicano[8]. A transferência da Diretoria Geral de Estatística, da Secretaria do Ministério do Império para o Ministério da Viação, também concorreu, até certo ponto, para dificultar esse trabalho. Em verdade, cinco anos depois de efetuado o recenseamento, ainda não havia sequer um estado onde as apurações do censo estivessem definitivamente concluídas. Só em 1900, às vésperas do II Censo Decenal da República, é que o Engenheiro Dr. Alfredo Eugênio de Almeida Maia, então Ministro da Viação, pôde anunciar a entrada no prelo da Diretoria Geral de Estatística do último volume da série de publicações do censo. Tal volume, segundo o programa estabelecido por aquela diretoria, devia conter a síntese dos resultados da pesquisa demográfica realizada nos primeiros dias do novo regime. De acordo com os resultados desse censo, a população do Brasil seria de 14 333 915 habitantes, dos quais 7 237 932 eram do sexo masculino e 7 095 983, do feminino.

[8] A revolta militar de 1893, por exemplo.

Como a Constituição da República, de 24 de fevereiro de 1891, determinava: "título primeiro. Seção I, Capitulo 1. Da Câmara dos Deputados. Art. 28. § 2º. Para esse fim mandará o Governo Federal proceder, desde já, ao Recenseamento da população da República, o qual será visto decenalmente", o Decreto nº 2.768, de 27 de dezembro de 1897, reformou a Diretoria Geral de Estatística no intuito de melhor adaptá-la à execução do recenseamento de 1900, para cujos trabalhos a Lei Orçamentária de 1898 consignou o necessário crédito. Em portaria de 20 de janeiro de 1900, foram aprovadas as instruções gerais que deviam reger o inquérito, assim como os modelos de instrumentos de coleta definidos pela Diretoria Geral de Estatística, de acordo com os votos do Instituto Internacional de Estatística, emitidos na sessão realizada em Berna, no ano de 1895.

Na data de 31 de dezembro de 1900, foi iniciado o censo em todo o território nacional. Os resultados das apurações na cidade do Rio de Janeiro, publicados em 29 de janeiro de 1901, suscitaram, porém, protesto geral, pela deficiência dos dados divulgados – críticas, aliás, formuladas por pessoas de elevado conceito social e de reconhecida autoridade em matéria de estatística – sendo nomeada uma comissão para estudar a matéria. O parecer conclusivo dessa comissão teve como conseqüência o cancelamento dos resultados do Censo de 1900, na parte referente ao Distrito Federal.

Apesar desse primeiro insucesso, que impôs a necessidade de se proceder a um novo inquérito na área territorial da Capital da República, prosseguiu a Diretoria Geral de Estatística na apuração dos resultados correspondentes às demais unidades políticas. Foi a seguinte a população apurada para o Brasil, de acordo como Censo de 1900, suprindo-se por estimativa a falta de informações relativas ao Distrito Federal: população total do País – 17 318 556 habitantes, sendo 8 831 002 homens e 8 487 554 mulheres.

Só em 1906, com base estabelecida nos Decretos de 6, 13 e 19 de junho desse ano, o Distrito Federal empreendeu novo inquérito[9], o qual foi concluído em fevereiro de 1907 e divulgado em volume especial elaborado pela Diretoria Geral de Estatística. O total da população do Distrito Federal era de 811 443 habitantes e, posteriormente, a Diretoria Geral publicou o montante de 17 438 434 para o Brasil.

O III Censo Decenal da República deveria realizar-se em 31 de dezembro de 1910, mas, por motivos conjunturais, cuja crítica minuciosa consta do relatório apresentado pelo Diretor Geral de Estatística ao Ministro da Agricultura, em fevereiro de 1915, deixou de ser levado a efeito na época prefixada, transferido que foi para 30 de junho do ano seguinte, sendo afinal definitivamente suspensos os respectivos trabalhos, após o dispêndio de avultadas verbas. Responde pelo malogro do Censo de 1910 o cerceamento da autonomia da Diretoria Geral de Estatística, por força das condições especiais em que se encontrava o País, então profundamente agitado por perturbações de ordem política.

Desde 1919, diligenciava a Direção Geral de Estatística firmar em bases seguras o empreendimento que seria a realização do IV Censo Decenal brasileiro. A certeza de que não lhe faltariam nem a imprescindível autonomia para uma ação eficiente contra os obstáculos a enfrentar, nem o prestígio indispensável para resistir, na defesa do interesse público, às perturbadoras influências que haviam prejudicado o êxito do censo anterior, foram, para o órgão diretor

[9] A iniciativa desse empreendimento coube à Prefeitura Municipal, que, na administração do Dr. Francisco Pereira Passos, resolveu levar a efeito o recenseamento do Rio de Janeiro.

do censo, os melhores estímulos à execução da tarefa que lhe confiara o Congresso Nacional, com base na Lei nº 4.017, de 9 de janeiro de 1920: a realização do Recenseamento Geral de 1920. Em 21 de janeiro desse mesmo ano, era aprovado o regulamento para a execução dessa Lei no Decreto nº 14.026. Entre as medidas preliminares tomadas pela Diretoria Geral de Estatística para a execução desse censo, citam-se: o envio de um emissário especial ao U.S. Census Bureau, dos Estados Unidos, para colher informações sobre os trabalhos preliminares do XIV Censo Americano e obter indicações precisas sobre como adquirir, por compra ou contrato, o equipamento mecânico utilizado na apuração de seus inquéritos. Cabia ainda a esse emissário averiguar as condições do mercado de papel, para fins de importação do material necessário. Decisão, também, muito importante, no âmbito da abrangência da pesquisa, há muito já delineada em suas linhas gerais, foi a de fazer a investigação estatística abranger não só a população, considerada no seu aspecto puramente demográfico, como também a situação econômica das várias localidades. Assim, a produção agrícola e industrial do País tornou-se objeto de meticuloso balanço.

Para a coleta das informações referentes às características da população, foi adotada a Lista de Família, em vez do Boletim Individual – de uso mais corrente nas operações censitárias realizadas em outros países – em razão das condições inerentes ao nosso meio. Além do modelo especial de questionário para os domicílios particulares, adotou a Diretoria de Estatística outro para as habitações coletivas. Nos questionários demográficos, foram feitas perguntas muito simples, de fácil resposta, sobre as principais características de cada um dos membros da família recenseada. A supressão do quesito relativo à cor foi explicada pelo fato de as respostas ocultarem parte da verdade, especialmente quanto aos mestiços. Outro quesito também suprimido nesse censo foi o referente à religião professada, porque o estudo estatístico das diversas confissões exorbitaria do caráter sintético que devem ter as investigações realizadas dentro do esquema de um recenseamento geral. No que diz respeito aos defeitos físicos, o Censo de 1920 cogitou, apenas, da cegueira e da surdo-mudez. O boletim utilizado no inquérito classificou os recenseados em três categorias: pessoas moradoras no domicílio e nele presentes na data do censo - 1º de setembro de 1920; pessoas moradoras no domicílio, mas dele ausentes nessa mesma data; pessoas não moradoras no domicílio, mas que ali passaram a noite de 31 de agosto de 1920. Isto permitiu que se separasse a população de fato da população de direito, pelo tratamento adequado desses grupos de recenseados.

Entre os principais instrumentos de coleta utilizados no Recenseamento Geral de 1920, figura a caderneta demográfica, apropriada ao registro pelos agentes recenseadores das informações referentes a cada prédio e das mais notáveis ocorrências verificadas durante os trabalhos de campo. Destinavam-se, basicamente, os registros nela constantes à organização da estatística predial.

Pelos relatórios dos delegados gerais, cujos registros foram confirmados pelas informações obtidas através dos delegados secionais, das comissões censitárias e até mesmo dos agentes recenseadores, o recenseamento da população e das indústrias foi, no conjunto, aceitável, aproximando-se seus resultados da realidade, observação infelizmente não extensiva ao censo agrícola e pastoril, dadas as sensíveis deficiências constatadas em sua realização, bastante comuns em pesquisa rural desse gênero e amplitude.

Os 18 179 recenseadores mobilizados nas operações desse censo levantaram uma população de 30 635 605 habitantes para o País em conjunto.

Dentro da periodicidade decenal dos censos brasileiros, prevista em lei, deveria realizar-se em 1930, por força do Decreto-lei nº 5.730, de 15 de outubro de 1929[10], o V Recenseamento Geral da população. Não obstante essa previsão, motivos, principalmente de ordem política, determinaram a sua não realização nesse ano.

Em 1940, transcorridos vinte anos da data da realização da última operação censitária, e após a radical transformação da estrutura dos serviços de estatística que se vinha processando nos últimos anos, realizou-se, a 1º de setembro de 1940, o V Recenseamento Geral do Brasil.

Pelo Decreto nº 24.609, de 6 de julho de 1934, era criado o Instituto Nacional de Estatística, entidade de natureza federativa, tendo por fim, mediante a progressiva articulação e cooperação das três ordens administrativas da Organização Política da República, bem como da iniciativa privada, promover e executar, ou orientar tecnicamente, em regime racionalizado, o levantamento de todas as estatísticas nacionais[11]. Mais tarde, esse instituto passou a denominar-se Conselho Nacional de Estatística e, dentro dos princípios de unidade nacional e de cooperação interadministrativa, passar a integrar, com o Conselho Nacional de Geografia,[12] o Instituto Brasileiro de Geografia e Estatística, por força do Decreto-lei nº 218, de 26 de janeiro de 1938.

Como igualmente previa o estatuto orgânico da instituição, o Decreto-lei nº 237, de 2 de fevereiro de 1938, completou o plano anteriormente delineado, criando a Comissão Censitária Nacional, órgão de direção superior do Serviço Nacional de Recenseamento, e que deveria, articuladamente com os demais órgãos do instituto, executar o Recenseamento Geral de 1940.

O Decreto-lei nº 311, de 2 de março de 1938, que instituiu o sistema de normas reguladoras do quadro territorial, permitiu a melhor delimitação das circunscrições do País, tendo em vista, principalmente, os trabalhos preparatórios do recenseamento geral da população e das atividades por esta desenvolvidas.

Em 21 de dezembro de 1938, era expedido o Decreto-lei nº 969, que dispôs sobre a realização dos recenseamentos gerais do Brasil e estabeleceu as normas sob as quais se processariam esses levantamentos. Em 15 de abril de 1940, foi esse Decreto-lei complementado pelo de nº 2.141, que regulamentou a execução do V Recenseamento Geral.

Planejados os instrumentos que deveriam abranger as investigações consideradas no esquema de seus sete censos[13] e instalada a rede nacional dos serviços censitários, realizou-se, em 1º de setembro de 1940, o Recenseamento Geral do Brasil, que correspondeu ao quinto censo da população; ao segundo censo da agricultura e da indústria e ao primeiro levantamento do comércio, dos transportes e comunicações, bem como dos serviços.

---

[10] Estatuto regulamentado pelo Decreto nº 18.994, de 19 de novembro de 1929.

[11] O decreto citado dotou o Instituto de uma organização provisória, deixando à futura Convenção Nacional de Estatística, em que estariam representados os Governos das três órbitas administrativas da República, a missão de fixar a definitiva estrutura do Sistema Estatístico Nacional.

[12] Antigo Conselho Brasileiro de Geografia, criado pelo Decreto nº 1.527, de 24 de março de 1937.

[13] O esquema geral de operações foi, resumidamente, o seguinte: censo demográfico (edificação, população); censo agrícola (exploração direta do solo, indústrias rurais); censo industrial (indústria de transformação, indústrias especiais); censo comercial (comércio de mercadorias, comércio de valores); censo dos transportes (transportes, comunicações); censo dos serviços (serviços pessoais, serviços coletivos); censo social (organizações urbanas, instituições sócio-culturais). Dentro do esquema geral deste censo, foram programados cinco inquéritos complementares.

Como já foi dito, o IV Censo de População, realizado em 1920, utilizou, na coleta de informações, a Lista de Família, que compreendia dois formulários, destinados ao recenseamento dos domicílios particulares e das habitações coletivas. Além de folhas suplementares desses instrumentos, foi igualmente usada a caderneta demográfica para fins já declarados e bem definidos.

O Censo Demográfico de 1940 substituiu aqueles instrumentos, sem alterar-lhes a essência, por um Boletim de Família, um Boletim Individual, a Lista de Domicílio Coletivo e a Caderneta do Agente Recenseador. A caderneta empregada neste censo, de uso exclusivo do agente recenseador, além de conter informações sumárias sobre o setor de operação a que se referia, facultou a obtenção de dados sobre os logradouros, os prédios e respectivos pavimentos, os domicílios e as pessoas neles recenseadas. Os resultados de apuração imediata e preliminar dos dados de população, contidos nessa caderneta, constaram da "Sinopse Preliminar do Censo Demográfico", que a Comissão Censitária Nacional fez divulgar em 1941.

Como instrumentos de coleta das informações de caráter eminentemente demográfico, foram usados o Boletim de Família e o Boletim Individual. Esses dois instrumentos, além das indicações referentes à localização e identificação do domicílio, particular ou coletivo, continham 45 quesitos a serem informados por todas as pessoas componentes das famílias censitárias ou pelo recenseado avulso. A lista de domicílio coletivo foi usada apenas para controlar a distribuição e resumir o movimento dos boletins relativos a tais domicílios.

Os caracteres biológicos e étnicos do recenseado – sexo, idade, cor – foram indagados em quatro quesitos. Seguia-se o quesito referente às enfermidades, cogitando-se apenas dos estados patológicos da cegueira, constitucionais ou adquiridos, e da surdo-mudez. Numa série de quesitos de interesse jurídico e social, indagou-se do estado civil e da descendência do recenseado. Outra série, de caráter político, diz respeito à naturalidade e à nacionalidade, bem como à língua falada no lar. A religião foi igualmente pesquisada. Seguiram-se várias indagações de natureza cultural, com o objetivo de focalizar o desenvolvimento da instrução do povo brasileiro. Dez quesitos referiam-se à ocupação do recenseado. Outras informações de ordem econômica e social, tais como inscrição em sindicatos, propriedade imobiliária, participação em instituições de previdência e seguros sociais, foram objeto de quesitos especiais. Os dois últimos quesitos dos boletins destinaram-se a verificar a qualidade eventual de "ausente temporário" ou "presente temporário" do recenseado no domicílio.

A apuração, elaboração e subseqüente divulgação dos resultados do Censo Demográfico de 1940 sofreram retardamento em razão de várias causas - deficiência quantitativa do equipamento mecânico disponível, demora na devolução do material de coleta preenchido, dificuldades de transporte, grande extensão territorial do País, etc. – sendo que, só em 1946, com a divulgação da Sinopse do Censo Demográfico, foram conhecidos os primeiros resultados definitivos desse censo. O Censo de 1940 registrou 41 165 289 habitantes no Brasil.

O Recenseamento Geral de 1950 constituiu a sexta operação censitária levada a efeito no País. Realizado de conformidade com a Lei nº 651, de 13 de março de 1949, que criou o Serviço Nacional de Recenseamento, em caráter transitório, na Secretaria Geral do extinto Conselho Nacional de Estatística, e

revigorou os princípios do Decreto-lei nº 969, de 21 de dezembro de 1938, que disciplinou a execução dessas operações, abrange os Censos Demográficos, Agrícola, Industrial, Comercial e dos Serviços, além de inquéritos especiais sobre transportes e comunicações[14].

A orientação geral dos trabalhos censitários, que competia à Comissão Censitária Nacional no Recenseamento Geral de 1940, foi atribuída à Junta Executiva Central, órgão deliberativo do Conselho Nacional de Estatística. As atividades que, em 1940, couberam às Delegacias Regionais, criadas para fins censitários, foram desempenhadas pelas Inspetorias Regionais de Estatística, que administravam, em cada Unidade da Federação, a Rede de Agências Municipais de Estatística. Com o fim específico de auxiliar os trabalhos de propaganda do recenseamento, foram instituídas Comissões Censitárias nas Unidades da Federação e nos municípios.

Os mapas municipais constituíram o elemento cartográfico mais importante utilizado nos trabalhos censitários e retratavam a divisão territorial fixada para o qüinqüênio 1944/1948. Na elaboração desses mapas, o Serviço Nacional de Recenseamento teve a colaboração dos órgãos geográficos regionais, e valeu-se, também, de elementos cartográficos pertencentes a entidades públicas e a organizações privadas. Para efeito de coleta, os municípios foram divididos em setores censitários, correspondendo o setor à menor unidade de trabalho.

O campo de investigação do Censo Demográfico de 1950 sofreu considerável redução em confronto com o de 1940. O número de quesitos baixou de 45 para 25, com a eliminação das perguntas referentes à cegueira, surdo-mudez, naturalidade dos pais do recenseado, data da fixação de residência no País, dos estrangeiros e brasileiros naturalizados, sindicalização, propriedade de imóveis, previdência social e seguros privados. Excluíram-se, ainda, dos boletins as indagações que visavam conhecer a idade do recenseado na data de nascimento do primeiro filho; o número de pessoas que falavam corretamente o Português; o de pessoas que recebiam instrução, a espécie, o grau dessa instrução e o local onde era ministrada; o de pessoas que se habilitaram em alguma arte ou ofício; e a espécie de remuneração na ocupação principal ou suplementar.

Contando com a experiência do censo anterior sobre a aceitabilidade do instrumento de coleta pelos informantes, o Serviço Nacional de Recenseamento testou seus modelos de boletim entre funcionários de várias repartições e em estabelecimentos industriais fora do perímetro urbano do Distrito Federal. Foram preenchidos 348 boletins com declarações referentes a 1 582 pessoas. Os quesitos se mostraram adequados, e as instruções, bem compreendidas.

A data de referência para o Censo Demográfico de 1950 foi fixada em 1º de julho. Foram recenseadas, em cada domicílio, as pessoas que aí passaram a noite de 30 de junho para 1º de julho, bem como os moradores efetivos, temporariamente ausentes, inclusive os menores internos em estabelecimentos de ensino de qualquer espécie.

[14] As normas gerais referentes e esses censos e inquéritos foram complementadas pelo Regulamento baixado pelo Decreto nº 26.914, de 20 de julho de 1949.

À semelhança do ocorrido no recenseamento precedente, o Censo Demográfico de 1950 empregou os seguintes instrumentos de coleta:

- Boletim de Família, destinado ao recenseamento das famílias censitárias;
- Boletim Individual, reservado ao registro da pessoa que ocupasse dependência de domicilio coletivo;
- Lista de Domicílio Coletivo, empregada para registrar a distribuição e o motivo dos boletins preenchidos em domicílios coletivos; e
- Caderneta do Recenseador, utilizada para o controle do trabalho de coleta e para o resumo de dados sobre os domicílios e as pessoas recenseadas.

Estudos do Gabinete Técnico do Serviço Nacional de Recenseamento, baseados nos resultados dos seis censos demográficos realizados no País, indicam a ocorrência de erro, por falta, no cômputo da população recenseada em 1900 e, por excesso, no registro da população recenseada em 1920. Ainda de acordo com esses estudos, nos resultados do Censo Demográfico de 1940, deixaram de ser computadas 17 913 pessoas recenseadas, cujas características não puderam ser apuradas, por extravio do material de coleta – 7 469 em parte do Município de Parintins, Estado do Amazonas, e 9 244 em parte do Município de Garça, Estado de São Paulo – enquanto nos resultados do Censo Demográfico de 1950 deixaram de ser incluídas 31 960 pessoas – 10 461 no Município de Nova Era, Minas Gerais, 7 588 no Município de Pirangi, São Paulo, e 13 911 no Município de Lapa, Paraná - por igual motivo[15].

Com a publicação da Sinopse Preliminar do Censo Demográfico, em março de 1951 e, bem assim, dos elementos básicos da investigação censitária, a partir de setembro de 1951, deu-se cumprimento à Lei nº 651, de 13 de março de 1949, que estabeleceu o prazo de dois anos para a divulgação dos resultados gerais e provisórios desse recenseamento. Em 1950, a população residente no país era de 51 941 767 habitantes.

Desde a segunda metade do século passado, o Brasil vem participando de Congressos Internacionais de Estatística e das seções periodicamente promovidas pelo Instituto Interamericano de Estatística – IASI –, visando ao desenvolvimento e à comparabilidade das estatísticas oficiais no âmbito das Américas. Assim, o Recenseamento de 1950, levado a efeito no País, integrou o Censo das Américas, de realização patrocinada por aquele Instituto, em atendimento a uma solicitação específica da Organização das Nações Unidas – ONU.

O Recenseamento Geral de 1960, sétima operação censitária realizada no País, seguiu, também, as diretrizes básicas estabelecidas por aqueles organismos internacionais, então incorporadas no Programa Mínimo do Censo das Américas[16].

Realizado em 1º de setembro, com base no Decreto-lei nº 969, de 21 de dezembro de 1938, que à época disciplinava a execução das operações censitárias, o Recenseamento de 1960 manteve o caráter inviolável e confidencial atribuído às informações prestadas para fins estatísticos e assegurou a não-utilização dessas informações, para fins de prova contra terceiros. O Serviço Nacional de Recenseamento, criado pelo Decreto nº 47.813, de 2 de março de 1960, em caráter tran-

[15] Os resultados do Censo de 1872 incluem 181 583 pessoas estimadas para as 32 paróquias de população não-recenseadas, enquanto nos resultados do Censo de 1900 computou-se para o Distrito Federal a população nele recenseada em 20 de setembro de 1906, em virtude de terem sido anulados os resultados daquele censo, correspondentes a essa Unidade da Federação.

[16] Esse Programa Mínimo visa assegurar a uniformidade de conceitos, do qual depende a comparabilidade dos resultados dos censos das nações americanas.

sitório, e integrado ao Instituto Brasileiro de Geografia e Estatística, foi o órgão executor do censo, auxiliado, nas atividades de coleta, pelas Inspetorias Regionais de Estatísticas, do Conselho Nacional de Estatística, que administravam, em cada Unidade da Federação, a rede de Agências Municipais de Estatística.

Os trabalhos de campo tiveram como base geográfica os mapas municipais especialmente elaborados para fins censitários, anteriormente ampliados, revistos e enriquecidos, com a colaboração dos agentes municipais de estatística. Para efeito de coleta, os municípios foram divididos em setores censitários, de tamanho variável com o número de unidades a serem recenseadas, formados por área territorial contínua situada num só quadro urbano, suburbano ou rural do mesmo distrito administrativo.

Foi realizado, em 17 de maio de 1959, o Censo de Brasília. Esse inquérito foi o segundo de caráter experimental levado a efeito pelo Núcleo de Planejamento Censitário. O primeiro foi executado na cidade de Ituiutaba, Minas Gerais. O Senhor Presidente da República manifestara interesse em conhecer, da forma mais ampla possível, a situação demográfica da nova capital. Assim, a escolha de Brasília atendeu a duplo objetivo.

O Censo Demográfico de 1960, conforme tradição dos censos brasileiros, pesquisou as características das pessoas, das famílias e dos domicílios. No que diz respeito às pessoas, investigou a situação no domicílio, condição de presença, sexo, condição no domicílio, idade, religião, cor, naturalidade, nacionalidade, migrações internas, instrução, estado conjugal, duração do casamento, fecundidade e algumas características econômicas. As tabulações referentes às famílias foram obtidas a partir de um tratamento adequado desses elementos.

Relativamente aos domicílios, o censo investigou o número de moradores, situação – urbana ou rural – da moradia, tipo de construção, condição de ocupação, aluguel mensal, forma de abastecimento de água, instalações sanitárias, combustível utilizado no fogão, número de cômodos, número de dormitórios e existência de iluminação elétrica, rádio, geladeira e televisão. A população aborígine só foi recenseada quando já aculturada, e os brasileiros residentes no exterior, ainda que em exercício de funções oficiais, não foram computados na população recenseada.

Utilizou este censo os seguintes instrumentos de coleta :

- Boletim Geral – C.D. 1, destinado ao recenseamento das pessoas residentes em domicílios particulares ou coletivos não-selecionados para cobrirem todo o campo de investigação da amostra;
- Boletim de Amostra – C.D. 2, utilizado para o recenseamento das pessoas residentes em domicílios particulares ou coletivos selecionados para a cobertura do campo de investigação da amostra, no qual se incluem as características dos domicílios particulares;
- Lista de Domicílios Coletivos – C.D. 3, empregada para controlar a distribuição dos boletins gerais e da amostra nos domicílios coletivos e permitir a seleção, nesses domicílios, das pessoas a serem recenseadas através do boletim de amostra;
- Caderneta do Recenseador – C.D. 6, destinada a reunir as folhas de coleta e resumir o trabalho do setor; e
- Folhas de Coleta – C.D. 7 e C.D. 8, utilizadas para o arrolamento das unidades do setor e permitir a seleção dos domicílios particulares a serem incluídos na amostra.

Em cada domicílio, foram recenseadas as pessoas que ali passaram a noite de 31 de agosto para 1º de setembro de 1960, bem como os moradores efetivos, temporariamente ausentes, inclusive os menores internados em estabelecimentos de ensino. Ao todo, o Censo de 1960 registrou uma população de 70 070 457 habitantes.

Pela primeira vez, foi empregada a técnica da amostragem em um censo brasileiro. Com o auxílio dessa técnica foram investigadas, através de nove quesitos, as características das pessoas, inclusive econômicas, e todas as informações referentes aos domicílios.

A amostra utilizada para isto foi constituída de aproximadamente 25% dos domicílios particulares e pessoas neles recenseadas, além de 25% dos grupos familiares ou componentes de grupos conviventes recenseados em domicílios coletivos.

A divulgação dos resultados preliminares do Censo Demográfico de 1960 foi iniciada em junho de 1961, com a apresentação da Sinopse Preliminar do Estado do Espírito Santo, e terminada em maio do ano seguinte, com a publicação dos resultados correspondentes ao País como um todo. Mas, só muito tempo depois, por razões de ordem operacional, foram conhecidos os resultados definitivos desse censo.

É oportuno fixar que o abandono do sistema convencional de apuração mecânica de dados, adotado pela repartição censitária nos últimos censos brasileiros, e a utilização de um computador de alto rendimento para esse fim específico – pela primeira vez usado no País, particularmente na operação de um censo – foram as causas diretas ou indiretas do atraso, medido em vários anos, verificado na operação dos resultados definitivos do Censo Demográfico de 1960.

Passados dez anos, a contar da data desse censo, realizou o Brasil, em 1970, o seu VIII Recenseamento Geral. Ainda aqui, as operações censitárias seguiram as mesmas diretrizes básicas recomendadas pelos organismos internacionais para o Censo das Américas, consubstanciadas no Programa Mínimo de investigação e tabulações, estabelecido para este Censo com a aprovação de todas as nações americanas.

Integrado no Censo das Américas – programa realizado sob o patrocínio do Instituto Interamericano de Estatística – IASI –, órgão pertencente à Organização dos Estados Americanos – OEA – o Recenseamento Geral de 1970 abrangeu os Censos Demográfico, Predial, Agropecuário, Industrial, Comercial e dos Serviços, além de inquéritos especiais sobre as instituições de crédito e seguradoras, bem como de produção e distribuição de energia elétrica.

Os estudos preliminares datam de 1967. No planejamento do Censo Demográfico de 1970, estava prevista uma pesquisa experimental que seria realizada em Contagem, Minas Gerais, para teste dos instrumentos de coleta das fases de processamento. Por falta de recursos financeiros, essa etapa foi suprimida.

A realização do Recenseamento Geral de 1970 obedeceu às determinações do Decreto-lei nº 369, de 19 de dezembro de 1968, regulamentado pelo Decreto nº 6.520, de 15 de maio de 1969, cujas disposições foram parcialmente alteradas pelo Decreto nº 65.697, de 12 de novembro do mesmo ano.

Nos termos do Decreto nº 64.520, foi atribuída ao IBGE, por intermédio do Departamento de Censos do Instituto Brasileiro de Estatística, a responsabilidade de planejar e executar o recenseamento, sob a assistência técnica da Comissão Censitária Nacional.

Dentro da estrutura geral dessa operação censitária, o Censo Demográfico de 1970, tendo como data de referência fixada o dia 1º de setembro desse ano, compreendeu a investigação das características individuais do recenseado – situação no domicílio, sexo, condição de presença, condição no domicílio, idade, religião, nacionalidade, naturalidade, migrações internas, instrução, estado conjugal, características econômicas e fecundidade; das famílias – composição e características de seus membros; e dos domicílios – número de moradores, situação, tipo de construção, condição de ocupação, aluguel mensal, tempo de residência, forma de abastecimento de água, existência e tipo de instalações sanitárias, combustível utilizado no fogão, número de cômodos, número de dormitórios e existência de iluminação elétrica, rádio, geladeira, televisão e automóvel.

Foram recenseadas, em cada domicílio, as pessoas que ali passaram a noite de 31 de agosto para 1º de setembro, bem como os moradores efetivos temporariamente ausentes, inclusive os menores de idade internados em estabelecimentos de ensino de qualquer espécie. Computaram-se, em tal Censo, os brasileiros em exercício de missão diplomática ou militar no exterior e os tripulantes e passageiros de navios brasileiros que se encontravam fora dos portos nacionais.

Não foram incluídos no Censo de 1970 os membros de representações diplomáticas ou militares dos países com os quais o Brasil mantém relações de amizade os tripulantes e passageiros de navios estrangeiros em trânsito pelos portos nacionais, bem como os aborígines que viviam em tribos, sem contato direto e permanente com a civilização. A população residente no país, registrada nesse censo, foi de 93 139 037 habitantes.

Colocado nos exatos moldes do censo anterior, o Censo Demográfico de 1970 utilizou os seguintes instrumentos de coleta:

- Boletim da Amostra – CD 1.01, utilizado para o recenseamento das pessoas residentes nos domicílios particulares ou coletivos selecionados para as indagações da amostra e para o levantamento das características dos domicílios particulares[17];
- Boletim da Não-amostra – CD 1.02, destinado ao recenseamento das pessoas residentes em domicílios particulares ou coletivos não-selecionadas para as indagações da amostra[18];
- Lista de Domicílios Coletivos – CD 1.03, empregada para controlar a distribuição dos boletins da amostra e da não-amostra nos domicílios coletivos, além de permitir a seleção, nesses domicílios, das pessoas a serem recenseadas através do primeiro desses boletins;
- Caderneta do Recenseador – CD 1.06, documento que apresenta o croqui e a descrição do setor censitário a ser coberto pelo agente e se destina ao registro do resumo das informações por ele coletadas bem como ao acondicionamento das folhas de coleta e das listas de domicílios coletivos; e
- Folhas de Coleta – CD 1.07 e CD 1.08, utilizadas para arrolar as unidades do setor e permitir a seleção dos domicílios particulares a serem incluídos na amostra.

[17] Correspondente ao boletim de amostra – C.D. 2, do Censo Demográfico 1960.
[18] Correspondente ao boletim geral – C.D. 1, do Censo Demográfico de 1960.

O esquema de amostragem aqui adotado foi idêntico ao do Censo de 1960. A amostra foi constituída por aproximadamente 25% dos domicílios particulares e pessoas neles recenseadas, além de 25% das famílias ou componentes de grupos conviventes recenseados em domicílios coletivos.

As tabulações dos itens investigados por amostragem foram obtidas usando-se o processo de estimativa por razão, no qual os fatores de expansão resultaram da divisão do total de pessoas no universo pelo total de pessoas na amostra. As razões obtidas em nível de município foram determinadas para 46 grupos de controle.

A fim de evitar ulteriores ajustamentos entre as tabulações da amostra, o censo utilizou "pesos" expressos em valores inteiros, imediatamente superiores ou inferiores à razão fracionária encontrada, na elaboração das estimativas. Assegurou-se, desta forma, que o produto desses "pesos" pelas unidades de amostra correspondesse sempre ao total das unidades do universo.

Em 1º de setembro de 1980, o IBGE realizou o IX Recenseamento Geral do País, obedecendo às determinações da Lei nº 5.878, de 11 de maio de 1973, e aos dispositivos do Decreto nº 74.084, de 20 de maio de 1974, bem como aos do Decreto nº 84.221, de 19 de novembro de 1979.

Para testar o plano elaborado em tal censo, o IBGE realizou, em 1979, um Censo Experimental no Município de Taubaté, São Paulo. Esse levantamento teve como objetivo colher elementos para o aperfeiçoamento de seus formulários, instruções, códigos, métodos e processos de coleta e apuração dos dados, visando à especificação do plano definitivo da pesquisa.

O Recenseamento Geral de 1980, que teve como slogan "O país que a gente conta", compreendeu os Censos Demográfico, Agropecuário, Industrial, Comercial e dos Serviços e, a exemplo dos levantamentos anteriores, integrou-se ao Censo das Américas, conforme orientação da ONU.

O Censo Demográfico de 1980 contou com 9 714 postos de coleta, instalados em 3 391 municípios, divididos em 141 553 setores. Contou também com avanços tecnológicos que permitiram apresentar inovações na sua realização e na divulgação dos resultados.

A primeira grande inovação foram o desenvolvimento e a utilização de um sistema informatizado de acompanhamento da coleta. Através de tal sistema, era possível conhecer, semanalmente, o número de setores concluídos e de pessoas neles recenseadas por sexo, bem como o tipo de questionário utilizado - amostra ou não-amostra - para acompanhar o andamento da coleta relativa a todo o Brasil, aqui incluídas as Unidades de Federação e os municípios.

Todas as Unidades de Federação enviavam as informações por telex. O material era examinado na sede do IBGE e encaminhado à Diretoria de Informática, que processava os dados contando com uma equipe de digitadores que perfuravam cartões. As informações eram colocadas no mainframe e dali se obtinha uma série de indicadores que permitiam acompanhar a evolução e detectar eventuais problemas no campo. Assim, a comunicação com os responsáveis nos estados era rápida, para que se verificasse por que em determinado lugar a coleta não obedecia a um ritmo normal, ou se investigasse a origem de outros problemas como, por exemplo, a fração amostral, a razão de sexos, a cobertura comparada com a projeção, a taxa de crescimento anual, entre outros.

Houve também, no Censo de 1980, uma novidade na divulgação, quando, pela primeira vez, os resultados saíram no mesmo ano de realização da pesquisa. Os resultados preliminares do Censo Demográfico de 1980 foram divulgados no dia 16 de dezembro desse mesmo ano. Isto foi uma inovação porque, tradicionalmente, não se tinha uma publicação preliminar. O que possibilitou tal divulgação foi a utilização do sistema de acompanhamento da coleta, que oferecia rapidez na análise dos dados. Quando os resultados foram fechados no sistema, o material estava praticamente pronto para ser divulgado, embora ainda sem passar pela etapa da crítica. Por isso, foram chamados resultados preliminares.

A publicação dos Resultados Preliminares do Censo de 1980 apresentou o total da população residente por sexo, além do total de não-moradores presentes e a taxa média geométrica de incremento anual da população residente. Os resultados foram divulgados para todo o Brasil, aí incluídas as grandes regiões, as Unidades da Federação, as regiões metropolitanas e os municípios.

No Censo Demográfico de 1980, foram recenseadas, em cada domicílio, as pessoas que ali passaram a noite de 31 de agosto para 1º de setembro de 1980, mesmo ausentes por período não-superior a 12 meses. Incluíram-se os brasileiros em missão diplomática ou militar fora do país e suas famílias, bem como os tripulantes de navios e aeronaves brasileiros no exterior. Não foram recenseados os estrangeiros em trânsito pelo nosso território, em navios ou aeronaves de diferentes origens na data do Censo, nem os aborígines que viviam em tribos, arredios ao contato. A população totalizou, nesse ano, 119 002 706 habitantes.

Foram pesquisadas as seguintes características relativas a:

- pessoas - situação do domicílio, sexo, condição de presença, condição no domicílio, idade, alfabetização, religião, cor, orfandade materna, estado conjugal, nacionalidade, naturalidade, migrações internas, freqüência à escola, nível de instrução, curso concluído, movimentos pendulares, características econômicas, fecundidade e mortalidade;
- composição das famílias; e
- domicílios - situação, número de moradores, espécie, tipo, condição de ocupação, material predominante nas paredes, piso e cobertura, forma de abastecimento de água, existência de escoadouro e uso das instalações sanitárias, aluguel ou prestação mensal, número de cômodos, cômodos com função de dormitório, existência e tipo de fogão, combustível usado para cozinhar, bem como existência de telefone, iluminação elétrica, rádio, geladeira, televisão e automóvel.

São os seguintes os instrumentos de coleta dignos de destaque, utilizados no referido censo:

- Boletim da Amostra – CD 1.01, formulário usado para o recenseamento dos domicílios da amostra e de pessoas neles residentes;
- Boletim da Não-amostra – CD 1.02, formulário usado para o recenseamento dos domicílios da não-amostra e de pessoas neles residentes;
- Lista de Domicílios Coletivos – CD 1.03, formulário usado para relacionar as famílias e as pessoas sós, recenseadas em cada domicílio coletivo, determinando em que unidades ou dependências seriam aplicados os boletins da amostra e da não-amostra;

- Caderneta do Recenseador – CD 1.06, documento com o mapa e a descrição do setor para orientação do recenseador da área, o registro do resumo das informações coletadas, o acondicionamento das folhas de coleta e as listas de domicílios coletivos;
- Folhas de Coleta – CD 1.07 e CD 1.08, listas de enumeração dos endereços, das espécies das unidades visitadas e dos nomes dos responsáveis pelos domicílios ocupados; instrumentos que determinaram os domicílios particulares da amostra e da não-amostra; registro do número de formulários utilizados e pessoas recenseadas;
- Boletim Resumo – CD 1.10, formulário para resumo dos trabalhos de coleta realizados em cada município; e
- Folha de Andamento da Coleta – CD 1.11, formulário para controlar o andamento dos trabalhos dos recenseadores durante a coleta de informações.

Assim como nos censos anteriores, a unidade de amostra do Censo de 1980 foi o domicílio. A amostra foi constituída por aproximadamente 25% dos domicílios particulares e pessoas neles recenseadas, além de 25% das famílias ou componentes de grupos conviventes recenseados em domicílios coletivos.

Os resultados do Censo Demográfico de 1980 foram classificados em dois grupos:

- resultados preliminares – englobavam os volumes de Resultados Preliminares e a Sinopse Preliminar, que apresentaram um primeiro conjunto de resultados referentes ao universo, cujos dados eram provenientes do boletim resumo, e os volumes das Tabulações Avançadas, que traziam resultados sobre as principais características investigadas, constantes no questionário da amostra, a partir de uma subamostra de aproximadamente 1% desse grupo; e
- resultados definitivos – continham os dados finais referentes a todas as características investigadas das pessoas, das famílias e dos domicílios.

Até a realização do censo seguinte, passaram-se 11 anos, não houve Censo em 1990. Para essa operação, o IBGE precisava contratar mais de 180 mil pessoas em caráter temporário. Entretanto, a autorização da contratação desse contingente foi sendo protelada, visto que o governo tinha a diretriz de reduzir o quadro de servidores públicos. Quando a contratação foi autorizada, em julho de 1990, não havia mais tempo hábil para se realizar o processo seletivo público para contratação dos recursos humanos destinados ao censo, cuja coleta de dados estava programada para se iniciar no mês de setembro. Por tal motivo, o IBGE decidiu realizar o censo em 1991.

O planejamento da pesquisa começou em 1987. Como nos censos anteriores, foi realizado um Censo Experimental, dessa vez em Limeira – SP -, no ano de 1989. Com o slogan "Ajude o Brasil a ter um bom censo", o XX Recenseamento Geral do Brasil teve início em 1º de setembro de 1991 e compreendeu apenas o Censo Demográfico. Sua realização obedeceu às determinações da Lei nº 8.184, de 10 de maio de 1991, e seguiu as recomendações da ONU e de outros organismos internacionais.

O Censo Demográfico de 1991 apresentou inovações em vários aspectos da pesquisa em relação aos levantamentos anteriores. A criação das Comissões Censitárias Municipais, para ajudar na mobilização da população e dar apoio à etapa de coleta de dados, foi uma das novidades. Compostas por representan-

tes de diversos segmentos da sociedade – governo, entidades de classe, associações comunitárias, instituições públicas e privadas – as comissões tinham como objetivo buscar junto a esses representantes e seus grupos, em nível municipal, apoio e parcerias para a realização da coleta, através do acompanhamento pela comunidade, do desenvolvimento dos trabalhos nos seus aspectos operacionais. As Comissões Censitárias colaboraram ainda na análise de mapas municipais, na instalação de postos de coleta, na divulgação do censo para a comunidade, entre outras tarefas.

Ainda dentro da mobilização, também pela primeira vez foi realizado o Projeto Escola no Censo de 1991. Tendo como público os alunos de estabelecimentos de ensino da rede pública, das turmas de alfabetização até a 8ª série, o projeto consistia em enviar material produzido por pedagogos e técnicos do IBGE para as escolas. Esse material continha sugestões de trabalhos para os professores desenvolverem em sala de aula. Desta forma, com as atividades propostas, os alunos estariam aprendendo mais sobre o censo e sobre o País, ao mesmo tempo em que poderiam ajudar na preparação para o recebimento dos recenseadores, através da mobilização de seus pais, amigos e outras pessoas do seu convívio.

A participação da sociedade em questões referentes ao censo também teve uma inovação com a criação da Comissão Consultiva, composta por estudiosos e especialistas em diversos temas ligados à população. O papel principal da Comissão Consultiva era ser um elo entre o IBGE e a sociedade. Com isso, buscou-se atender à demanda de informações dos diversos setores, contando com a assessoria desses especialistas, que opinavam nas decisões relativas ao conteúdo dos questionários, à amostra, aos métodos de apuração, entre outros aspectos.

Na parte técnica da pesquisa, pode-se citar como inovações a introdução de novos quesitos no Questionário Básico – rendimento e grau de escolaridade do chefe do domicílio - além da ampliação dos quesitos referentes à caracterização do domicílio. Note-se também que, pela primeira vez nos censos contemporâneos, pesquisou-se a deficiência física e mental, através do Questionário da Amostra.

É justamente quanto à dimensão da amostra que se teve a principal inovação do Censo de 1991. Após muitos estudos e discussões, concluiu-se pela sua redução, que nos Censos de 1960 a 1980 era de 25% dos domicílios. Assim, em 1991, foram definidas duas frações amostrais; 10% dos domicílios, nos municípios com mais de 15 000 habitantes, e 20% nos demais, o que representava cerca de 4 milhões de questionários.

Outra inovação foi a entrada de dados descentralizada nas Unidades da Federação, o que permitiu mais agilidade nessa etapa. Os sistemas de administração e de pagamento utilizaram, pela primeira vez, recursos da informática. Na fase de divulgação, das informações, a utilização de disquetes também foi uma novidade. Dessa forma, os resultados do Censo de 1991 estavam disponíveis em volumes impressos, fitas magnéticas, sistemas informatizados de consulta e disquetes.

No censo em questão, foram recenseadas, em cada domicílio, as pessoas residentes no território nacional na data de referência, inclusive os brasileiros em missão diplomática ou militar fora do país e suas famílias, bem como os tripulantes de navios e aeronaves brasileiros no exterior. Não foram recenseados os estrangeiros em trânsito pelo nosso território, em navios ou aeronaves de diferentes origens, na data do censo, nem os aborígines que viviam em tribos, arredios ao contato, conservando hábitos primitivos. O resultado revelou que o Brasil tinha, naquela época, 146 825 475 habitantes.

Foram investigadas as características das pessoas, das famílias e dos domicílios, assim distribuídas:

- características das pessoas – sexo, condição no domicílio, idade, cor ou raça, religião ou culto, deficiência física ou mental, orfandade materna, nacionalidade, naturalidade, migrações internas e internacionais, instrução, nupcialidade, fecundidade, mortalidade e características econômicas;
- composição das famílias; e
- características dos domicílios – localização; material predominante nas paredes e na cobertura; forma de abastecimento de água; tipo de escoadouro e uso de instalação sanitária; condição de ocupação; aluguel mensal; combustível usado para cozinhar; número de cômodos, dormitórios e banheiros; destino do lixo; número de linhas telefônicas e automóveis particulares; existência e propriedade de automóvel para trabalho; existência e forma de iluminação; número de aparelhos de televisão em cores e existência de filtro de água, rádio, televisão em preto e branco, freezer, geladeira, máquina de lavar roupa e aspirador de pó.

Os principais instrumentos de coleta utilizados no referido censo foram os seguintes:

- Questionário Básico – CD 1.01, formulário usado para o recenseamento dos domicílios da não-amostra e das pessoas neles residentes;
- Questionário da Amostra – CD 1.02, formulário usado para o recenseamento dos domicílios da amostra e das pessoas neles residentes;
- Lista de Domicílio Coletivo – CD 1.03, formulário usado para relacionar as famílias e as pessoas sós, recenseadas em cada domicílio coletivo, determinando em que unidades ou dependências seriam aplicados os questionários básico e da amostra;
- Caderneta do Setor – CD 1.06, documento com o mapa e a descrição do setor para orientação do recenseador da área, o registro do resumo das informações coletadas, o acondicionamento das folhas de coleta e as listas de domicílios coletivos;
- Folha de Coleta – CD 1.07, lista de enumeração dos endereços, das espécies das unidades visitadas e dos nomes dos responsáveis pelos domicílios ocupados; servia também para determinar os domicílios particulares da amostra e da não-amostra, registrando o número de formulários utilizados e pessoas recenseadas;
- Boletim Resumo – CD 1.10, formulário para resumo dos trabalhos de coleta realizados em cada município; e
- Folha de Andamento da Coleta – CD 1.11, formulário para controlar o andamento dos trabalhos dos recenseadores durante a coleta de informações.

Assim como no censo anterior, os resultados do Censo Demográfico de 1991 foram classificados em dois grupos:

- resultados preliminares – englobavam os volumes de Resultados Preliminares e a Sinopse Preliminar, que apresentaram um primeiro conjunto de resultados do universo, cujos dados eram provenientes do Boletim Resumo; e
- resultados definitivos – continham os dados finais referentes a todas as características investigadas das pessoas, das famílias e dos domicílios.

O Censo Demográfico 2000, no qual foi retomada a realização em anos terminados em zero, foi o décimo primeiro recenseamento realizado no País, revelando as características da população na entrada do novo século.

Uma mudança em relação aos censos anteriores foi a antecipação da data de referência, que era 1º de setembro; no Censo 2000, essa data passou a ser 1º de agosto.

No âmbito internacional, o Censo 2000 significou a consolidação dos laços estatísticos entre os países do Mercosul Ampliado, que inclui os membros do Mercosul – Brasil, Argentina, Paraguai e Uruguai – além de Bolívia e Chile, através do Projeto Censo Comum do Mercosul. Iniciado em 1997, o projeto permitiu a padronização de conceitos e classificações, visando homogeneizar e fortalecer os sistemas estatísticos nacionais e, finalmente, criar uma base de dados comum aos censos dos seis países.

Como teste final dos métodos, procedimentos, equipamentos e materiais para o Censo 2000, o IBGE realizou o Censo Experimental em 1999, nas cidades de Marília – SP - e Bonito – PA – para definir o plano definitivo da pesquisa.

De uma forma geral, pode-se dizer que o Censo 2000 foi marcado pela inovação tecnológica, começando pelo sistema que permitiu o acompanhamento da coleta, que operou através de um site na Internet, passando pela captura de dados, onde ocorreu a digitalização dos questionários e o reconhecimento ótico de caracteres, chegando à automação dos processos de codificação, crítica e tabulação dos dados.

1.2 Legislação

O Censo Demográfico 2000 foi realizado com respaldo na legislação geral que regula os levantamentos das estatísticas nacionais, dispõe sobre a obrigatoriedade da prestação de informações, protege o caráter confidencial das informações coletadas e prevê a divulgação e a entrega de resultados ao público. A execução do Censo 2000 obedeceu ainda à legislação que disciplina e respalda as ações de caráter administrativo-gerencial, quais sejam: a contratação de pessoal temporário no setor público, a aquisição de bens e serviços e as dotações orçamentárias.

A seguir, são apresentadas as leis e decretos federais relacionados, direta ou indiretamente, com a operação censitária.

Decreto-lei nº 969, de 21 de dezembro de 1938 – dispõe sobre os Recenseamentos Gerais do Brasil. Nos termos de seu Artigo 12, § 3º, foram concedidas as Gratificações por Serviço Extraordinário aos servidores do IBGE que desempenharam funções de coordenação e supervisão durante a operação do Censo 2000.

Lei nº 5.534, de 14 de novembro de 1968 – dispõe sobre a obrigatoriedade da prestação de informações estatísticas e dá outras providências. Essa lei foi modificada pela Lei nº 5.878, de 11 de maio de 1973, regulamentada pelo Decreto nº 73.177, de 20 de novembro de 1973.

Lei nº 5.878, de 11 de maio de 1973 – dispõe sobre a Fundação Instituto Brasileiro de Geografia e Estatística – IBGE, e dá outras providências. Seu Artigo 2º define: "constitui objetivo básico do IBGE assegurar informações e estudos de natureza Estatística, Geográfica, Cartográfica e Demográfica necessári-

as ao conhecimento da realidade física, econômica e social do País, visando especialmente ao Planejamento Econômico e Social e à Segurança Nacional". O inciso I do seu Artigo 3º estabelece que, para a consecução desse objetivo básico, a Instituição atuará, principalmente, na área das estatísticas primárias, contínuas e censitárias.

Decreto nº 73.177, de 20 de novembro de 1973 – regulamenta a Lei nº 5.534, de 14 de novembro de 1968, modificada pela Lei nº 5.878, de 11 de maio de 1973, que dispõe sobre a obrigatoriedade de prestação de informações ao Plano Nacional de Estatísticas Básicas e ao Plano Geral de Informações Estatísticas e Geográficas. Esse decreto estabelece quem está obrigado a prestar as informações solicitadas pelo IBGE e o caráter sigiloso que resguarda as informações prestadas.

O Decreto nº 74.084, de 20 de maio de 1974 – aprova o Plano Geral de Informações Estatísticas e Geográficas, assegura, em seu Artigo 8º, a proteção, pelo sigilo, das informações previstas no plano e, em seu Artigo 9º, prevê a divulgação de informações e dados, bem como sua entrega ao público através de anuários, relatórios, sinopses e publicações especializadas.

Lei nº 7.853, de 24 de outubro de 1989 – dispõe sobre o apoio às pessoas portadoras de deficiência, sua integração social e dá outras providências. Seu Artigo 2º, inciso III, item d, trata da reserva de mercado de trabalho, em favor das pessoas portadoras de deficiência, nas entidades da Administração Pública e do setor privado.

Lei nº 8.184, de 10 de maio de 1991 – dispõe sobre a periodicidade dos censos demográficos e dos censos econômicos e dá outras providências. Seu Artigo 1º determina: " A periodicidade dos Censos Demográficos e dos Censos Econômicos, realizados pela Fundação Instituto Brasileiro de Geografia e Estatística (IBGE), será fixada por ato do Poder Executivo, não podendo exceder a dez anos a dos Censos Demográficos e a cinco anos a dos Censos Econômicos". A autorização do Poder Executivo para a realização do Censo 2000 deu-se pela inclusão do Projeto Censo Demográfico 2000 no Plano Plurianual 1999 – 2002 do Governo Federal, a partir de proposta apresentada pelo IBGE ao Ministério do Planejamento, Orçamento e Gestão.

Lei nº 8.666, de 21 de junho de 1993 – regulamenta o Artigo 37, inciso XXI, da Constituição Federal, institui normas para licitações e contratos da Administração Pública e dá outras providências.

Lei nº 8.745, de 9 de dezembro de 1993 – dispõe sobre a contratação de pessoal por tempo determinado, para atender à necessidade temporária de excepcional interesse público, nos termos do inciso IX do Artigo 37 da Constituição Federal, e dá outras providências. O item III de seu Artigo 2º trata das contratações para "a realização de recenseamentos e outras pesquisas de natureza estatística efetuadas pela Fundação Instituto Brasileiro de Geografia e Estatística – IBGE" .

Decreto nº 2.424, de 17 de dezembro de 1997 – autoriza excepcionalmente a contratação temporária de recursos humanos para os órgãos e entidades do Poder Executivo Federal, que especifica, entre eles, o Instituto Brasileiro de Geografia e Estatística.

Medida Provisória nº 1.748-40, de 8 de abril de 1999 – altera artigos da Lei nº 8.745, de 9 de dezembro de 1993, que dispõe sobre a contratação por tempo determinado para atender à necessidade temporária de excepcional interesse

publico, e dá outras providências. Esta medida provisória incluiu o inciso III do Artigo 2º, que estendeu a autorização para contratação de pessoal temporário.

Lei nº 9.849/99, de 26 de outubro de 1999 – altera artigos da Lei nº 8.745, de 9 de dezembro de 1993, que dispõe sobre a contratação por tempo determinado para atender à necessidade temporária de excepcional interesse público, e dá outras providências.

Decreto nº 3.272, de 3 de dezembro de 1999 – aprova novo Estatuto e o Quadro Demonstrativo dos Cargos em Comissão e Funções Gratificadas da Fundação Instituto Brasileiro de Geografia e Estatística, e dá outras providências. Estabelece seu Artigo 29: "As atividades censitárias serão custeadas por dotações específicas consignadas ao IBGE no Orçamento da União, conforme disposto no Art. 15 da Lei nº 5.878, de 11 de maio de 1973".

Decreto nº 3.298, de 20 de dezembro de 1999 – regulamenta a Lei nº 7.853, de 24 de outubro de 1989, e dispõe sobre a política nacional para a integração da pessoa portadora de deficiência. Seus dispositivos foram aplicados na contratação de pessoal temporário, implicando reserva de vagas para pessoas portadoras de deficiência física nos processos seletivos.

1.3 Periodicidade e data de referência

No que toca à periodicidade, os censos demográficos têm sido decenais em quase todo o mundo, e o foram no Brasil desde 1890, com as exceções de 1910, 1930 e 1990. A importância de se respeitar o intervalo decenal é inquestionável entre os especialistas, pois a recomposição de uma população nacional com base em censos posteriores ou anteriores não é tarefa trivial, nem em países com população estável.

No Brasil, a quebra do intervalo decenal, em função do adiamento do Censo Demográfico de 1990 para 1991, exigiu ajustes anuais para calcular indicadores, dificultando a interpretação de seus resultados, de modo que certas dúvidas só serão definitivamente resolvidas com a utilização das informações do Censo 2000.

Quanto à data de referência, o melhor período para a realização de censos demográficos é aquele no qual o movimento da população é o mínimo possível, ou seja, durante o qual a grande maioria das pessoas está presente em seus domicílios. Dois outros aspectos são as condições climáticas durante a coleta de dados e a relação da programação do censo com o ano fiscal, uma vez que operações censitárias requerem liberação de recursos e contratações, ambas em grande escala.

Assim, a escolha da data para os censos nacionais é uma questão importante, que tem impactos na precisão e comparabilidade dos dados apurados. Levando-se em conta a diversidade climática, o adiantamento do calendário das eleições municipais do mês de novembro para outubro e os ganhos oriundos da integração e padronização com os censos do Mercosul, decidiu-se adotar como data de referência o dia 1º de agosto de 2000.

Com isso, a comparabilidade com as informações dos censos passados – datas de referência em 1º de setembro – não ficava prejudicada, eram reduzidos os efeitos da proximidade das eleições sobre a coleta da pesquisa, além de garantido o tempo suficiente para seu término dentro de um mesmo ano fiscal.

Referências

ALMEIDA, C. M. de. *Atlas do Imperio do Brazil*: comprehendendo as respectivas divisões administrativas ecclesiasticas, eleitoraes e judiciarias... destinado á instrucção publica no Imperio com especialidade á alumnos do Imperial Collegio de Pedro II. Rio de Janeiro: Lith. do Instituto Philomathico, 1868. Dedicado ao Imperador D. Pedro II.

BRASIL. Regulamento nº 120, de 31 de janeiro de 1842. Regula a execução da parte policial e criminal da Lei nº 261, de 3 de dezembro de 1841. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 31, 1842. Col. 1.

_____. Lei nº 387, de 19 de agosto de 1846. Regula a maneira de proceder às eleições de senadores, deputados, membros das assembléias provinciais, juízes de paz, e câmaras municipais. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 13, 1846. Col. 1.

_____. Lei nº 586, de 6 de setembro de 1850. Manda reger no exercício de 1851-1852 a Lei do Orçamento nº 555, de 15 de junho do corrente ano. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 279, 1850. Col. 1.

_____. Decreto nº 797, de 18 de junho de 1851. Manda executar o regulamento para a organização do Censo Geral do Império. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 161, 1851. Col. 1.

_____. Decreto nº 798, de 18 de junho de 1851. Manda executar o regulamento do registro dos nascimentos e óbitos. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 168, 1851. Col. 1.

_____. Decreto nº 907, de 29 de janeiro de 1852. Suspende a execução dos regulamentos para a organização do Censo Geral do Império e para o registro dos nascimentos e óbitos. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 19, 1852. Col. 1.

_____. Decreto nº 1.829, de 9 setembro de 1870. Sanciona o Decreto da Assembléia Geral que manda proceder o recenseamento da população do Império. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 89, 1870.

_____. Decreto nº 4.676, de 14 de janeiro de 1871. Cria na Corte do Império uma Diretoria Geral de Estatística, em virtude de autorização concedida pelo Artigo 2º da Lei nº 1.829 de 9 de setembro de 1870, e manda executar o respectivo regulamento. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 35, 1871. Col. 1.

_____. Decreto nº 4.856, de 30 de dezembro de 1871. Manda proceder, em execução do Artigo 1º da Lei nº 1.829 de 9 de setembro de 1870, ao primeiro recenseamento da população do Império. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 786, 1871. Col. 1.

_____. Decreto nº 2.940, de 31 de outubro de 1879. Fixa a despesa e orça a receita geral do Império para os exercícios de 1879-1880 e 1880-1881, e dá outras providências. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 104, 1879. Col. 1.

_____. Decreto nº 113-d, de 2 de janeiro de 1890. Restaura e reorganiza a Diretoria Geral de Estatística, criada pelo Artigo 2º da Lei nº 1.829 de 9 de setembro de 1870, e manda proceder ao segundo recenseamento da população dos Estados Unidos do Brasil. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 2, 1890. Col. 1.

BRASIL. Decreto nº 331, de 12 de abril de 1890. Dá novo regulamento à Diretoria Geral de Estatística. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 3, p. 600, 1890. Col. 1.

_____. Decreto nº 659, de 12 de agosto de 1890. Manda observar as instruções para o segundo recenseamento da população dos Estados Unidos do Brasil. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 8, p. 1843, 1890. Col. 1.

_____. Constituição (1891). *Constituição da República dos Estados Unidos do Brasil*. Rio de Janeiro: Imprensa Nacional, 1891. 39 p.

_____. Decreto nº 2.768, de 27 de dezembro de 1897. Altera o pessoal e dá nova distribuição aos serviços a cargo da Diretoria Geral de Estatística. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 967, 1897. Col. 1.

_____. Lei nº 4.017, de 9 de janeiro de 1920. Autoriza o Governo a mandar proceder, no dia 1 de setembro de 1920, ao recenseamento geral da população do Brasil e dá outras providências. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 174, 1920. Col. 1.

_____. Decreto nº 14.026, de 21 de janeiro de 1920. Aprova o regulamento para execução da Lei nº 4.017, de 9 de janeiro de 1920. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 456, 1920. Col. 1.

_____. Decreto-lei nº 5.730, de 15 de outubro de 1929. Autoriza a proceder ao recenseamento geral da República em 1 de setembro de 1930, e dá outras providências. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 1, p. 73, 1929. Col. 1.

_____. Decreto nº 18.994, de 19 de novembro de 1929. Dá regulamento ao Decreto Legislativo nº 573, de 15 de outubro de 1929, que autoriza proceder ao recenseamento geral da República em setembro de 1930. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 4, p. 202, 1929. Col. 1.

_____. Decreto nº 24.609, de 6 de julho de 1934. Cria o Instituto Nacional de Estatística e fixa disposições orgânicas para a execução e desenvolvimento dos serviços estatísticos. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 4, p. 524, 1934. Col. 1.

_____. Decreto nº 1.527, de 24 de março de 1937. Institui o Conselho Brasileiro de Geografia, anexo ao Instituto Nacional de Estatística, autoriza a sua adesão a União Geográfica Internacional e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 1 abr. 1937. p. 7187. Col. 2.

_____. Decreto-lei nº 218, de 26 de janeiro de 1938. Muda o nome do Instituto Nacional de Estatística e o do Conselho Brasileiro de Geografia. *Coleção de Leis [da] República Federativa do Brasil*, Brasília. DF, v. 1, p. 59, 1938. Col. 1.

_____. Decreto-lei nº 237, de 2 de fevereiro de 1938. Regula o início dos trabalhos do Recenseamento Geral da República em 1940, e dá outras providências. *Coleção de Leis [da] República Federativa do Brasil,* Brasília, DF, v. 1, p. 68, 1938. Col. 1.

_____. Decreto-lei nº 311, de 2 de março de 1938. Dispõe sobre a divisão territorial do país, e dá outras providências. *Coleção de Leis [da] República Federativa do Brasil,* Brasília, DF, v. 1, p. 438, 1938. Col. 1.

BRASIL. Decreto-lei nº 969, de 21 de dezembro de 1938. Dispõe sobre os recenseamentos gerais do Brasil. *Coleção de Leis [da] República Federativa do Brasil,* Brasília, DF, v. 4, p. 315, 1938. Col. 1.

_____. Decreto-lei nº 2.141, de 15 de abril de 1940. Regulamenta a execução do Recenseamento Geral de 1940, nos termos do Decreto-lei nº 969, de 21 dezembro de 1938. *Coleção de Leis [da] República Federativa do Brasil*, Brasília, DF, v. 3, p. 37, 1940. Col. 1.

_____. Lei nº 651, de 13 de março de 1949. Dispõe sobre o VI Recenseamento Geral do Brasil. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 23 mar. 1949. P. 420. Col. 1.

_____. Decreto nº 26.914, de 20 de julho de 1949. Aprova o regulamento do VI Recenseamento Geral do Brasil. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 23 jul. 1949. p. 10514. Col. 1. Retificado no Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF, 20 jul. 1949. p. 11465. Col. 4.

_____. Decreto nº 47.813, de 2 de março de 1960. Institui o Serviço Nacional de Recenseamento e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 2 de mar. 1960. p. 3457. Col. 1.

_____. Lei nº 5.534, de 14 de novembro de 1968. Dispõe sobre a obrigatoriedade de prestação de informações estatísticas e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 18 nov. 1968. p. 9985. Col. 2. Retificada no Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF, 20 nov. 1968. p. 10081. Col. 1.

_____. Decreto-lei nº 369, de 19 de dezembro de 1968. Dispõe sobre a realização do VIII Recenseamento Geral do Brasil de 1970. *Diário Oficial [da] República Federativa do Brasil,* Poder Executivo, Brasília, DF, 20 dez. 1970. p. 11019. Col. 4.

_____. Decreto nº 64.520, de 15 de maio de 1969. Aprova o regulamento do VIII Recenseamento Geral do Brasil. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 19 maio 1949. p. 4234. Col. 1. Retificado no Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF, 27 maio 1969. p. 4491. Col. 1.

_____. Lei nº 5.878, de 11 de maio de 1973. Dispõe sobre a Fundação Instituto Brasileiro de Geografia e Estatística - IBGE, e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 15 maio 1973. p. 4697. Col. 4.

_____. Decreto nº 73.177, de 20 de novembro de 1973. Regulamenta a Lei nº 5.534, de 14 de novembro de 1968, modificada pela Lei nº 5.878 de 11 de maio de 1973, que dispõe sobre a obrigatoriedade da prestação de informações necessárias ao Plano Nacional de Estatísticas Básicas e ao Plano Geral de Informações Estatísticas e Geográficas. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 21 nov. 1973. p. 11899. Col. 1.

BRASIL. Decreto nº 74.084, de 20 de maio de 1974. Aprova o Plano Geral de Informações Estatísticas e Geográficas, e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 21 maio 1974. p. 5773. Col. 2. Retificado no Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF, 5 ago. 1974. p. 8775. Col. 1.

_____. Decreto nº 84.221, de 19 de novembro de 1979. Dispõe sobre a realização do IX Recenseamento Geral do Brasil. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 20 nov. 1979. p. 17227. Col. 1.

_____. Lei nº 7.853, de 24 de outubro de 1989. Dispõe sobre o apoio às pessoas portadoras de deficiência, sua integração social, sobre a Coordenadoria Nacional para Integração da Pessoa Portadora de Deficiência - CORDE, institui a tutela jurisdicional de interesses coletivos ou difusos dessas pessoas, disciplina a atuação do Ministério Público, define crimes, e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 25 out. 1989. p. 1920. Col. 1.

_____. Lei nº 8.184, de 10 de maio de 1991. Dispõe sobre a periodicidade dos Censos Demográficos e dos Censos Econômicos e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 13 maio 1991. p. 8893. Col. 1.

_____. Lei nº 8.666, de 21 de junho de 1993. Regulamenta o artigo 37, inciso XXI, da Constituição Federal, institui normas para licitações e contratos da Administração Pública e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 22 jun. 1993. p. 8269. Col. 1. Retificada no Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF, 2 jul. 2003. p. 1. Col. 3.

_____. Lei nº 8.745, de 9 de dezembro de 1993. Dispõe sobre a contratação por tempo determinado para atender a necessidade temporária de excepcional interesse público, nos termos do inciso IX do artigo 37 da Constituição Federal, e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 10 dez. 1993. p. 18937. Col. 2.

BRASIL. Decreto nº 2.424, de 17 de dezembro de 1997. Autoriza excepcionalmente a contratação temporária de recursos humanos para os órgãos e entidades do Poder Executivo Federal que especifica. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 18 dez. 1997. p. 30225. Col. 2.

_____. Medida provisória nº 1.748-40, de 8 de abril de 1999. Altera os artigos 2, 3, 4, 5, 6, 7 e 9 da Lei nº 8.745, de 9 de dezembro de 1993, que dispõe sobre a contratação por tempo determinado para atender a necessidade temporária de excepcional interesse público, e dá outra providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 8 abr. 1999. p. 11. Col. 1.

_____. Lei nº 9.849, de 26 de outubro de 1999. Altera os artigos 2º, 3º, 4º, 5º, 6º, 7º e 9º da Lei nº 8.745, de 9 de dezembro de 1993, que dispõe sobre a contratação por tempo determinado para atender a necessidade temporária de excepcional interesse público, e dá outra providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 27 out. 1999. Seção Extra, p. 4. Col. 2.

BRASIL. Decreto nº 3.272, de 3 de dezembro de 1999. Aprova o estatuto e o quadro demonstrativo dos cargos em comissão e funções gratificadas da Fundação Instituto Brasileiro de Geografia e Estatística – IBGE, e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 6 dez. 1999. p. 4. Col. 1.

_____. Decreto nº 3.298, de 20 de dezembro de 1999. Regulamenta a Lei nº 7.853, de 24 de outubro de 1989, dispõe sobre a política nacional para a integração da pessoa portadora de deficiência, consolida as normas de proteção, e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 21 dez. 1999. p. 10. Col. 1.

CENSO DEMOGRÁFICO 1991: características gerais da população e instrução: resultados da amostra. Brasil. Rio de Janeiro: IBGE, n. 1, 1991.

_____: resultados do universo relativos às características da população e dos domicílios. Brasil. Rio de Janeiro: IBGE, n. 1, 1991.

COMITÊ do censo 2000: problemas relacionados com o adiamento do levantamento do Censo Demográfico de 1990 para o ano de 1991. Rio de Janeiro: IBGE, Diretoria de Pesquisas, [19—]. 4 p.

CONTAGEM da população 1996. Rio de Janeiro: IBGE, 1997. v.1: Resultados relativos a sexo da população e situação da unidade domiciliar. 728 p.

MEMORIA estatistica do Imperio do Brazil. *Revista Trimestral do Instituto Histórico e Geográfico Brasileiro*, Rio de Janeiro, t. 58, pt. 1, n. 91, p. 91-99, 1895.

METODOLOGIA do censo demográfico de 1980. Rio de Janeiro: IBGE, 1983. 478 p. (Série relatórios metodológicos , v. 4).

PRINCIPLES and recommendations for population and housing censuses. New York: United Nations, 1998.

RELATÓRIO apresentado ao Ministro e Secretario D'Estado dos Negocios do Imperio pela Comissão encarregada da direcção dos trabalhos do arrolamento da população do Municipio da Côrte... Rio de Janeiro: Typ. Perseverança, 1871. 36 p. tab.

SILVA, J. N. de S. e. *Investigações sobre os recenseamentos da população geral do Imperio e de cada provincia de per si tentados desde os tempos coloniaes até hoje*. Rio de Janeiro: Typ. Perseverança, 1870. 167 f.

2 Planejamento

O planejamento do Censo Demográfico 2000 teve início em 1997, quando foram realizadas as primeiras reuniões técnicas na Diretoria de Pesquisas – DPE – do IBGE com o objetivo de elaborar uma proposta inicial de conteúdo da pesquisa. Na mesma época, foram constituídas duas comissões: a Comissão de Planejamento e Organização Geral do Censo 2000 – CPO – de âmbito interno, com a atribuição de ditar as diretrizes gerais para o planejamento, a organização e a execução da operação censitária; e a Comissão Consultiva do Censo Demográfico 2000, composta por especialistas da comunidade técnico-científica, usuários das informações censitárias, convidados a assessorar a Diretoria de Pesquisas no planejamento do censo.

A complexidade e a amplitude de uma operação censitária exigem planejamento detalhado, exaustivo e cuidadoso, uma vez que seu preparo depende do encadeamento perfeito de inúmeras atividades de caráter técnico, administrativo-gerencial e de logística, e sua execução determina limites estreitos para a correção de eventuais falhas no decorrer de qualquer etapa. Assim, o planejamento adequado é condição essencial para a garantia do sucesso da operação. Na busca dessa condição e com o firme propósito de alcançar os melhores níveis de qualidade e transparência em todas as etapas da operação, o Censo Demográfico 2000 teve seu planejamento pautado na incorporação de novas tecnologias, de forma a colocar à disposição das equipes envolvidas e incorporar aos processos de trabalho as facilidades e os benefícios da evolução tecnológica.

Buscou-se, ainda, propiciar ampla divulgação do recenseamento propriamente dito, permitindo aos usuários dos dados censitários e à sociedade, de um modo geral, o acompanhamento de

cada etapa da operação e o acesso aos resultados em curto prazo e por meio das mais modernas mídias de comunicação e disseminação de informações.

No que diz respeito às inovações tecnológicas, algumas delas merecem destaque pelos impactos positivos que propiciaram ao desenvolvimento das diversas etapas da operação, à qualidade dos dados e à apuração e divulgação de resultados. Vale ressaltar a construção de bases digitais, que representou a mudança de patamar tecnológico – do formato convencional em papel para arquivo digital – de mapas em escalas cadastrais para as áreas urbanas e de mapas em escalas topográficas para as áreas totais dos municípios. A busca desse novo patamar foi, desde o princípio, colocada como fundamental para melhorar a qualidade da cobertura do território e adequar-se às novas mídias e tecnologias de disseminação de informações. Embora o ideal fosse abranger todos os municípios, sabia-se, desde o início do planejamento, que isso não seria viável, dadas as restrições de tempo e de recursos humanos e financeiros. Optou-se, então, por dar prioridade aos municípios com maior número de habitantes, de forma a garantir que as áreas que comportam a maior parcela de população do País fossem atendidas por um mapeamento de melhor qualidade, assegurando-lhes melhores níveis de cobertura e reduzindo problemas inerentes à coleta nas áreas de grande concentração populacional. Assim, o processo de conversão digital alcançou 100% dos mapas das áreas urbanas e 82% dos mapas municipais e gerou malhas digitais de todos os recortes político-administrativos legais – municípios, distritos e subdistritos – e operacionais – os setores censitários.

Outro aspecto que mereceu especial atenção na etapa de planejamento foi a necessidade de implementação de sistemas automatizados que dessem conta do controle da etapa de coleta de dados, oferecendo informações atualizadas e rápidas sobre a evolução dos trabalhos de campo nos vários níveis gerenciais, de acordo com os recortes territoriais de interesse – setor censitário, distrito, município, áreas regionais de coordenação, estados e o País. Entre os objetivos principais de tal demanda estavam o controle dos prazos da coleta de dados, a detecção eficiente de problemas na coleta, permitindo a adoção imediata de ações corretivas, a automação completa do pagamento dos recenseadores, buscando agilizar e eliminar erros inerentes aos processos manuais, bem como a consolidação dos dados tão logo fosse encerrada a coleta, visando à divulgação de resultados preliminares ainda no mesmo ano. Com base nessas premissas e demandas, foi desenvolvido o Sistema de Indicadores Gerenciais da Coleta – SIGC – que, disponibilizado na Internet, atendeu a essas e outras demandas surgidas ao longo do período de planejamento, preparo e execução da operação censitária e funcionou como um eficaz instrumento de gerenciamento da etapa de coleta de dados, inclusive no pagamento dos Recenseadores. Posteriormente, o SIGC forneceu informações iniciais para o Sistema de Indicadores Gerenciais da Produção – SIGP – semelhante ao SIGC, porém implementado durante os trabalhos de captura e apuração dos dados.

A mudança tecnológica de maior impacto no Censo 2000, no entanto, deu-se na forma de capturar os dados dos questionários. Desde o Censo de 1960, o IBGE vinha fazendo uso da digitação na conversão dos dados para o meio eletrônico. A Contagem da População 1996 usou a leitura ótica de marcas, tecnologia já considerada em obsolescência no final da década de 90 e com sérias restrições quanto ao uso em grande escala. Nesse contexto, no ano de 1997, quando o tema captura de dados foi posto em discussão no âmbito

do planejamento do Censo 2000, avaliaram-se todas as alternativas disponíveis – digitação, leitura de marcas com leitoras óticas e reconhecimento ótico de imagens digitalizadas com uso de escâneres. Foi examinada, inclusive, uma alternativa mista – uso de leitoras óticas no Questionário Básico e de escâneres no Questionário da Amostra – na tentativa de aproveitar conhecimento já adquirido e equipamentos disponíveis na Instituição. Para tanto, foram desenvolvidos diversos estudos, testes e avaliações. Entre eles, vale destacar a realização de duas provas piloto e as visitas técnicas feitas aos institutos oficiais de estatística da Argentina e dos Estados Unidos, que também testavam o reconhecimento ótico de imagens, respectivamente em prova-piloto e censo experimental. Respaldado, então, em conclusões favoráveis a respeito da alta taxa de confiabilidade nos *softwares* de reconhecimento e na baixa taxa de erros, o IBGE decidiu pela adoção do reconhecimento ótico de marcas e caracteres com uso de escâneres no Censo 2000, levando em conta também a satisfatória relação custo/benefício, assim considerada em vista dos ganhos de qualidade e homogeneidade, pela menor interferência humana e economia de tempo. Isso asseguraria a divulgação de todos os resultados em prazos menores do que os praticados em censos anteriores, e ainda a possibilidade de absorção dessa tecnologia para aplicação em outras pesquisas e trabalhos desenvolvidos na Instituição.

A partir de tal decisão, foram necessários outros estudos, testes e grandes reformulações nos modelos anteriormente usados dos questionários e acessórios de coleta, desde o formato e as características do papel até os cuidados com o manuseio e preenchimento dos formulários – caligrafia, grafite, borracha, prancheta. Com isso, houve também a necessidade de adaptação no transporte e armazenamento do material, alcançando também o treinamento dos Recenseadores.

Além dos estudos que nortearam as escolhas de novas tecnologias, o planejamento também abrangeu outros importantes itens, quais sejam: a avaliação das operações censitárias de 1996, que se encontrava em sua fase final quando foram dados os primeiros passos do planejamento do Censo 2000. Essa avaliação gerou uma série de recomendações para a nova operação; o envolvimento dos participantes externos, a integração com os países do Mercosul e, por fim, o orçamento – peça fundamental para assegurar o equilíbrio entre a otimização e a disponibilização adequada de recursos financeiros – itens estes que serão abordados no decorrer do presente capítulo.

2.1 Avaliação das operações censitárias de 1996 e recomendações para o Censo 2000

Durante o ano de 1998, com a contribuição das diversas áreas temáticas do IBGE, foi desenvolvido um trabalho junto às Unidades Regionais – URs - para conhecer, daqueles que mais diretamente participaram das atividades de campo das operações censitárias de 1996, as experiências e a avaliação quanto aos aspectos que mais contribuíram ou prejudicaram o alcance dos resultados do projeto. O objetivo final era, então, o enriquecimento das áreas envolvidas para melhoria da qualidade do planejamento e desenvolvimento do Censo 2000.

Após análise dos problemas levantados e das sugestões enviadas pelos representantes das diversas áreas do IBGE, foram feitos relatórios sínteses com as avaliações das unidades da sede da Instituição, a partir das avaliações re-

gionais e cuja consolidação gerou um relatório final com as sinalizações dos pontos importantes para as decisões que envolveriam o planejamento do Censo 2000. As principais recomendações são listadas a seguir.

2.1.1 Processo seletivo

- Fazer o teste seletivo separado para o Agente Censitário Municipal – ACM – e Agente Censitário Supervisor – ACS[1].
- Contratar para o quadro de pessoal temporário com antecedência de, no mínimo, dois meses do início da atividade.
- Exigir 2º grau de escolaridade para os cargos de ACM e ACS.
- Exigir 1º grau de escolaridade para o cargo de Recenseador.
- Diminuir o período de inscrições.
- Considerar como ideal a relação de um setor para cada Recenseador.
- Incluir no edital o valor médio da remuneração do Recenseador e a especificação clara das atribuições de cada função.
- Eliminar as restrições legais que impeçam a contratação de ex-censitários, aposentados e servidores públicos com disponibilidade de horário.
- Ficar a cargo do IBGE a seleção para ACM, ACS e pessoal de apoio.
- Ficar a cargo de empresa especializada a seleção para Recenseador.
- Dar divulgação do processo seletivo, pois isso é mais importante do que um período extenso de inscrição.
- Fazer a divulgação do processo seletivo, de forma intensa e antecipada, abrangendo obrigatoriamente os municípios pequenos e os distritos e povoados isolados.
- Fazer processo seletivo diversificado e em épocas distintas, permitindo a seleção de pessoal mais adequado ao perfil das tarefas e com menor índice de desistência, muito embora possa acarretar um aumento substancial de custos e da carga de trabalho nas URs; devem ser, então, avaliadas as vantagens e desvantagens dessa opção.
- Achar um nível de especificação ideal para o edital, pois um detalhamento exagerado pode gerar falta de mobilidade na época de contratação.
- Estabelecer alternativas para efetuar o pagamento das inscrições: cheque correio, vale postal, ou ordem de pagamento.
- Continuar definindo o valor da taxa de inscrição relacionado ao valor do salário e de mercado. É uma forma de triagem de candidatos.
- Definir, no ato de inscrição, para diminuir a evasão, o município ou agência onde o candidato concorrerá, muito embora a prova não deva ser realizada no município onde há vaga, pela complexidade da operação.
- Estabelecer processos seletivos específicos, porque o exercício de cada cargo deve exigir diferentes níveis de escolaridade.
- Selecionar os melhores classificados pois, caso contrário, seriam necessários métodos complexos de seleção.
- Fazer a prova diferenciada por região.
- Divulgar as inscrições por município para facilitar o preenchimento das vagas.

[1] Denominado, também, nesta metodologia, como Supervisor.

- Dar treinamento a um número de candidatos aprovados superior ao número das vagas, para suprir as eventuais desistências e reprovações no treinamento.
- Planejar o quantitativo de pessoal contratado para não atrasar a coleta.
- Melhorar a atividade gerencial na área de apuração.

2.1.2 Treinamento

- Aumentar a quantidade de material, considerando margem maior de reserva técnica.
- Distribuir o material com mais antecedência.
- Antecipar a disponibilização dos recursos financeiros para pagamentos de ajudas de custo e locação de instalações.
- Evitar os locais inadequados, locando melhores instalações para a realização das aulas.
- Adotar modernas técnicas de instrução – teleconferências, vídeo, multimídia e outros – onde não se elimine a presença do instrutor, que passará a atuar como monitor, visando garantir a homogeneidade das instruções e maior abrangência ao evento, assegurando assim que, em todas as etapas, o treinamento ministrado seja o mesmo.
- Regionalizar as etapas habitualmente centralizadas, visando atender às características regionais.
- Garantir a supervisão de técnicos da Instituição em todas as fases do treinamento, em especial a de Recenseador.
- Estabelecer um tempo maior – pelo menos uma semana – entre as etapas do treinamento, de tal forma que seja possível o exame e avaliação das entrevistas a título de teste.

2.1.3 Pagamento

- Garantir equipamentos dedicados exclusivamente às tarefas administrativas.
- Alocar técnicos da Diretoria Executiva – DE – da Diretoria de Informática – DI – e da Coordenação de Acompanhamento e Controle Operacional dos Censos – COC – em equipes para atendimento às URs; esse pessoal não poderá estar envolvido com as tarefas rotineiras do IBGE, nem com outros processos relativos ao censo demográfico. Sua formação terá que ser feita com um prazo confortável a fim de que todos dominem os sistemas adotados e tenham condições de ministrar treinamentos, dar suporte e atendimentos localizados.
- Prever treinamentos intensificados tanto para os operadores diretos do sistema quanto para o pessoal de menor nível da estrutura estadual, a fim de que todas as dúvidas sejam esgotadas antes da efetiva execução dos pagamentos.
- Contratar pessoas para cargos de apoio administrativo, tendo como pré-requisito a escolaridade em nível de 2º grau, além de experiência em assuntos relacionados à área administrativa, a fim de atuarem nos postos de coleta, agências, Divisões de Pesquisas – DIPEQs e Departamentos Regionais – DEREs – como responsáveis pela parte administrativo-financeira.

- Estabelecer procedimentos de pagamento atrelados aos dados da Caderneta do Setor, após serem conferidos pelo Supervisor, como fontes de alimentação do sistema de pagamento.
- Desenvolver sistema para ser utilizado em rede ligada ao Rio de Janeiro, com intuito de controlar e apurar os processos de concessão dos suprimentos de fundos, com orientações específicas para censos, emitidas pela área financeira a fim de garantir a padronização dos procedimentos.
- Desenvolvimento de sistema específico para pagamento do pessoal censitário em geral, com a fundamental característica de funcionamento em rede, ligando às agências, DIPEQ, DERE e Rio de Janeiro.
- Estabelecer um período de 90 dias, após o término dos testes definitivos do sistema de pagamento, para fins de elaboração do manual, disponibilização de material para audiovisual, local para treinamento, etc.
- Fazer uma simulação para testes do sistema antes do Censo 2000.

2.1.4 Recursos financeiros

- Definir um fluxo de crédito orçamentário e financeiro que se adapte às necessidades das unidades gestoras.
- Estabelecer maior integração entre IBGE e o Ministério de Planejamento e Orçamento – MPO – para que o fluxo financeiro seja disponível de acordo com a necessidade dos trabalhos.
- Adequar o plano de contas às atividades do censo.
- Definir o fluxo de informações gerenciais.
- Definir as unidades gestoras e seus aspectos administrativos, avaliando a disponibilidade de pessoal com sua respectiva formação técnica, de equipamentos e de prestadores de serviços, bancos, principalmente.
- Melhorar a relação do IBGE com os bancos, a fim de aprimorar o fluxo dos serviços bancários contratados.

2.1.5 Recursos materiais

- As Unidades Regionais deverão identificar, com antecedência mínima de seis meses, a necessidade de espaço próprio, que poderá ser cedido, compartilhado ou locado.
- Formalizar contrato, através do termo, de acordo com as normas e legislações vigentes, com antecedência, de no mínimo, três meses, na hipótese de locação.
- Trocar apenas correspondência entre o cedente e o cessionário para a hipótese de cessão de instalações, tendo em vista o pouco tempo de permanência nesses locais.
- Implementar a busca de espaços cedidos e/ou compartilhados, visando minimizar custos.
- Definir os critérios de compra - centralizado e descentralizado – levando-se em conta o tipo e a quantidade do material a ser adquirido, bem como o fato de que geralmente, as aquisições efetivadas por região costumam ser mais ágeis e vantajosas para a administração.
- Observar os prazos mínimos considerados suficientes para a realização de aquisição por licitação, de forma que a carta convite compreenda dez dias, e a tomada de preços, corresponda ao período de 45 dias.

- Agilizar os procedimentos de contratação de serviços definindo um plano de trabalho a ser aprovado e formalizado por Resoluções da Presidência do IBGE – R.PR – para que a Coordenação de Recursos Materiais – CRM – providencie as minutas de convite/edital e contrato, obtendo exame e aprovação da Procuradoria Geral – PGE – de um modelo único.
- Treinar e capacitar servidores para desempenho das atividade administrativas, considerando-se que as DIPEQs não dispõem de estrutura administrativa para desenvolvimento das atividades inerentes à CRM.
- Fornecer o conjunto de normas e legislações específicas, juntamente com o treinamento administrativo.

2.1.6 Base cartográfica

- Instituir como atividade permanente a atualização da base.
- Criar estrutura nas URs para a efetiva implementação da referida atividade, incluindo capacitação técnica.
- Envolver a rede de agências em um trabalho permanente de incorporação e base de todas as atualizações efetuadas pelas pesquisas contínuas.
- Envolver a supervisão da base nas DIPEQs, relativamente ao processo de treinamento dos recenseadores.
- Estabelecer, por parte da Diretoria de Geociências – DGC – junto a rede de coleta, um cronograma anual de trabalho, nos moldes do que é feito pela DPE.
- Formalizar parcerias com outras Instituições produtoras e usuárias de material cartográfico. Para tanto, há que se proceder a uma descentralização das atividades da DGC e a uma maior autonomia e recursos para as DIPEQs.

2.1.7 Comissões censitárias

- Conscientizar os representantes da sociedade da sua importância e criá-las com bastante antecedência em relação à data de início do censo, de forma que as comissões estaduais possam ser criadas em 1999 e as comissões municipais a partir de abril de 2000.
- Criar comissões nos municípios com mais de 10.000 habitantes.
- Garantir, sempre, a presença de um funcionário do IBGE como presidente da comissão.
- Fazer com que o IBGE arque com os custos de instalação, manutenção e infra-estrutura dos postos de coleta.

2.1.8 Informática

- Realizar experiências para melhor estabelecer a distribuição do trabalho entre as fases de processamento, antes do seu início, permitindo assim obter parâmetros de performance.
- Determinar a quantidade de lotes máximos de trabalho, por etapa, para não sobrecarregar a máquina.
- Estabelecer horários adequados para a realização de *back-ups* e rotinas de manutenção de bancos.
- Informatizar as agências.

2.1.9 Leitora ótica

- Estabelecer contratos de manutenção dos equipamentos em melhores condições, devendo haver assistência técnica em cada pólo de apuração.
- Melhorar o controle de qualidade do material de coleta, antes da utilização.
- Realizar experiências com novos equipamentos antes de adotá-los, a fim de medir melhor a sua performance.

2.1.10 Coleta

- Definir previamente o conteúdo dos questionários para não atrasar as fases de diagramação, impressão, teste nos equipamentos de leitura e, finalmente, a distribuição.
- Estabelecer previamente o quantitativo de questionários para não atrasar o andamento dos trabalhos.
- Estabelecer, em média, de um para cinco, a relação Supervisor/Recenseador.
- Desobrigar o Supervisor de exercer funções administrativas.

2.1.11 Apuração

- Planejar melhor as áreas de apuração, alugando espaço fora ou melhorando as condições das DIPEQs. Atentar para as condições de refrigeração e acústica.
- Definir previamente o conteúdo dos questionários para não atrasar a elaboração do sistema de crítica e imputação.

2.1.12 Divulgação dos resultados

- Sugerir a promoção de cursos de curta duração para as Unidades Regionais - chefias e pessoal com função gerencial – para que esses gerentes sejam dotados de conhecimento básico sobre as principais componentes da população em seus estados – história, migração, padrões de fecundidade, etc. – capacitando-os não somente para a fase de divulgação dos resultados mas, principalmente, para acompanhar com elementos técnicos a evolução da coleta, intervindo no processo quando necessário, durante a execução e não apenas ao final, o que acarreta custos e desgastes infinitamente superiores.
- Divulgar os dados com a ressalva "sujeito a modificações" e analisados em nível local/estadual.
- Divulgar os dados, através do projeto escola, seminários e palestras.
- Sugerir que as DIPEQs, agências e comissão censitária recebam os resultados, através de disquetes e de forma extra-oficial.
- Adquirir equipamentos adequados para todas as agências.
- Desenvolver *folders* para estudantes e pesquisadores, sendo que a distribuição deve ocorrer em ambiente de grande movimentação.

2.1.13 Publicidade

- Distribuir nota com o pronunciamento do Sr. Presidente da República e carta do Presidente do IBGE para todos os veículos de comunicação.
- Estabelecer convênios com outros órgãos para emissão de mensagens sobre o censo.

- Usar "carro de som" na publicidade do censo.
- Desenvolver o projeto escola e instituir a semana do censo nas escolas.
- Começar a campanha publicitária com, no mínimo, 60 dias de antecedência.
- Divulgar para todos os contratados os trabalhos realizados pelo IBGE.
- Distribuir bonés, livros didáticos e colocar propaganda na contra-capa dos cadernos escolares do Ministério da Educação e Cultura – MEC.

Essas recomendações foram analisadas e avaliadas. Muitas delas vieram auxiliar, efetivamente, na busca de maior qualidade na execução do Censo Demográfico 2000 – ver documento *Avaliação dos Censos 96: relatório final* (1998), incluído no CD-ROM que acompanha esta publicação.

2.2 Participantes externos

O planejamento do Censo Demográfico 2000 mobilizou todas as unidades do IBGE e envolveu consultas às comunidades de especialistas e a outros usuários das informações censitárias, além de parcerias com órgãos estaduais, municipais e internacionais.

2.2.1 Comissão Consultiva

A Comissão Consultiva do Censo Demográfico 2000 foi criada em 1997, sendo presidida pelo Diretor da Diretoria de Pesquisas e constituída por oito especialistas da comunidade técnico-científica, usuários das informações censitárias, a saber:

Charles Curt Mueller – Universidade de Brasília – UnB;

Elisa Maria da Conceição Pereira Reis – Instituto de Filosofia e Ciências Sociais – IFCS/UFRJ;

Elza Salvatori Berquó – Centro Brasileiro de Análise e Planejamento – CEBRAP/SP e Comissão Nacional de População – CNPD/DF;

José Alberto Magno de Carvalho – Centro de Desenvolvimento e Planejamento Regional – CEDEPLAR/MG;

Ricardo Paes e Barros – Instituto de Pesquisas Econômicas Aplicadas – IPEA/RJ;

Tânia Bacelar de Araújo – Universidade Federal de Pernambuco – UFPE; e

Wilton de Oliveira Bussab – Fundação Getulio Vargas – FGV/SP.

Coube à Comissão Consultiva assessorar a Diretoria de Pesquisas no planejamento do censo demográfico, em especial nas definições de conteúdo dos questionários, critérios de seleção e de estimação da amostra, nos métodos de apuração dos dados e na definição do plano de divulgação dos resultados.

A Comissão Consultiva foi convocada e reuniu-se com regularidade, desde 1997 até o ano de 2003, acompanhando a coleta dos dados e a avaliação e análise dos resultados, tanto do Censo Experimental como do Censo Demográfico 2000.

Discutiram-se no âmbito da comissão, as principais questões referentes ao Censo Demográfico 2000, passando pela formulação das perguntas dos questionários e suas categorias de investigação, como também pela validação da retirada e da inclusão de alguns quesitos.

Os principais temas discutidos foram os referentes às questões sobre cor ou raça, características dos domicílios, relações de parentesco, características econômicas da população, migração e deslocamento, fecundidade, mortalidade, bem como aspectos da logística da operação, entre outros.

2.2.2 Especialistas e usuários

Na preparação de grandes pesquisas, como a de um censo de população, é necessário fazer consultas não só ao segmento acadêmico das universidades e centros de pesquisa, mas também aos pesquisadores em demografia, educação e saúde, trabalho e renda, pobreza e desigualdades sociais, entre outras áreas de conhecimento. Esses especialistas constituem uma parcela importante dos usuários das informações do censo.

Para o Censo Demográfico 2000, foram enviadas 8 000 cartas aos membros do cadastro de usuários de informações do IBGE. As respostas recebidas foram aproximadamente 3% do total. Essa parcela tinha a seguinte composição: prefeituras e órgãos municipais 50,7%; ministérios, secretarias de estado e outros órgãos executivos estaduais e federais, 14,8%; conselhos, órgãos de pesquisa e fundações, 15,3%; universidades, 10,5%; e outros usuários, 8,6%.

Outro segmento importante dos usuários e interlocutores técnicos do IBGE é o dos órgãos regionais de estatística, responsáveis, em geral, pela área de planejamento de governos estaduais e prefeituras. Podem ser citados, dentre outros:

- no Rio de Janeiro, a Empresa Municipal de Informática S/A – IPLAN-Rio e a Fundação Centro de Informações e Dados do Rio de Janeiro – Fundação CIDE;
- em São Paulo, a Fundação Sistema Estadual de Análise de Dados – Fundação SEADE – e a Empresa Paulista de Planejamento Metropolitano S/A – EMPLASA;
- em Minas Gerais, o Centro de Desenvolvimento e Planejamento Regional – CEDEPLAR – a Empresa de Informática e Informação do Município de Belo Horizonte – PRODABEL – a Fundação João Pinheiro e o Instituto de Pesquisas Econômicas, Administrativas e Contábeis de Minas Gerais – IPEAD;
- no Rio Grande do Sul, a Fundação Estadual de Planejamento Metropolitano e Regional – METROPLAN – e a Fundação de Economia e Estatística Siegfried Emanuel Heuser – FEE;
- no Paraná, o Instituto Paranaense de Desenvolvimento Econômico e Social – IPARDES;
- no Ceará, o Instituto de Pesquisa e Informação do Estado do Ceará – IPLANCE;
- na Bahia, a Superintendência de Estudos Econômicos e Sociais da Bahia – SEI;
- no Distrito Federal, a Companhia de Desenvolvimento do Planalto Central –CODEPLAN; e
- no Pará; o IDESP - Instituto do Desenvolvimento Econômico e Social do Pará.

Um terceiro segmento de parceiros é o que colabora na atualização da Base Operacional e Geográfica – BOG. Trata-se também de órgãos estaduais e municipais, além de concessionárias de serviços públicos e empresas privadas,

que usam, produzem e/ou arquivam mapas e cadastros do território. Sem a colaboração desses parceiros, a atualização da base territorial seria impossível e graças a eles o Censo 2000 veio a contar com uma grande parcela da base operacional em meio digital.

A partir das discussões no âmbito do Comitê do Censo Demográfico 2000[2] da Diretoria de Pesquisas, surgiu a necessidade da colaboração de especialistas temáticos na formulação dos planos de crítica e imputação. Para tanto, acordaram-se diversas apresentações sobre o sistema de crítica e imputação de dados a serem utilizados no Censo, com o intuito de disseminar a metodologia e incorporar as sugestões dos especialistas. Essas apresentações foram feitas na Universidade Federal de Minas Gerais – UFMG – na Universidade de Campinas – UNICAMP – e na Escola Nacional de Ciências Estatísticas – ENCE – do IBGE, e resultaram em insumos importantes para o sistema de apuração do Censo Demográfico 2000.

2.3 Integração com os países do Mercosul, Chile e Bolívia

Desde 1991, tem havido propostas e ações no sentido de tornar efetivo o processo de integração entre os países do Bloco Mercosul. A partir de 1992, foram firmados compromissos vinculados às atividades estatísticas e agendadas tarefas conjuntas, para atingir a meta de homogeneizar e fortalecer os sistemas estatísticos nacionais.

Com relação à produção e à disponibilização de dados, os países membros vêm se aprimorando em decorrência do atendimento a demandas crescentes dos diversos usuários, que acessam freqüentemente a informação através das redes de informática.

Durante a década de 1990, as datas dos censos dos países membros do Mercosul ampliado foram distintas: Argentina e Brasil realizaram a pesquisa em 1991; Paraguai, Bolívia e Chile, em 1992; e Uruguai, em 1996.

Através do estudo comparado dos instrumentos de coleta dos censos de cada país, chegou-se à conclusão de que as definições conceituais e a operacionalização das variáveis não eram semelhantes. Surgiram também divergências na implementação e aplicação prática da pesquisa no campo, havendo, portanto, muita dificuldade em se fazer a comparação estatística entre os resultados dos censos. As restrições impostas por peculiaridades nacionais deram margem à necessidade de se ter um levantamento de informações básicas, que permitissem construir um panorama internacional.

Em virtude da ONU ter recomendado que os censos fossem realizados em anos terminados por zero, e o do ano 2000 ter marcado a mudança de milênio, os censos próximos desse ano constituíram um ponto de confluência dos esforços dos países do Bloco Mercosul para um trabalho em comum.

2.3.1 Projeto Censo Comum no Mercosul

O projeto de compatibilização dos censos de população dos países do Mercosul ampliado iniciou-se em maio de 1997, com a primeira oficina de discussão que teve como finalidade averiguar a situação dos países com rela-

[2] Para conhecer sua composição e forma de atuação, ver Seção 3.1.2.

ção às pesquisas dos censos anteriores, conhecer o planejamento das atividades relativas ao censo do ano 2000, aprovar conteúdos mínimos para integrar o próximo censo, bem como estabelecer mecanismos de intercâmbio para efetuar reuniões e oficinas.

A partir dessa primeira reunião, ficaram registrados compromissos e uma proposta de elaboração de um projeto de trabalho comum. Deste modo, os órgãos de estatística dos países do Mercosul realizaram, desde então, trabalhos e acordos em conjunto, com o propósito de estabelecer níveis de comparabilidade, entre variáveis, conceitos, definições e classificações aplicados nos censos de cada um dos países envolvidos. Com a aprovação do Projeto Censo Comum no Mercosul, em dezembro de 1997, foram iniciadas as seguintes tarefas:

- otimizar os recursos metodológicos, humanos e tecnológicos de cada país para a realização dos censos, unindo toda a experiência adquirida;
- viabilizar meios de apoio institucional e financeiro, para formalizar a iniciativa em uma data comum a todos os países;
- obter informações homogêneas a partir da inclusão de conteúdos mínimos com definições comuns, tanto conceituais quanto operacionais;
- processar e difundir os dados dos censos no formato adequado;
- facilitar o intercâmbio de experiências através da cooperação horizontal;
- potencializar o apoio de caráter internacional, bem como a assistência técnica.

Em conseqüência, foi gerado um espaço de gestão, integração e organização, para ratificar procedimentos das estatísticas nacionais e o planejamento dos censos.

O Projeto Censo Comum no Mercosul previu a criação de grupos técnicos, a realização de provas-piloto conjuntas para, em pequena escala, pôr em prática os questionários dos censos, de acordo não só com os conteúdos mínimos estabelecidos, as metodologias de capacitação, a organização das atividades necessárias para orientar o planejamento dos censos, mas também com o processamento da informação, a geração de dados e a divulgação comum dos acordos via Internet.

2.3.2 Planejamento do Censo Comum

Nas discussões entre os representantes dos países avaliaram-se as experiências dos últimos censos e os planos futuros, mas os principais tópicos tratados referiam-se à comparabilidade entre os dados produzidos em cada país. Como não seria possível harmonizar todas as variáveis de todos os censos, optou-se por selecionar um conjunto de variáveis comuns, e para estas, compatibilizar as definições e conceitos. Por outro lado, considerou-se importante tentar alcançar uma homogeneização operacional, como também garantir a compatibilidade dos sistemas de classificação e codificação. O objetivo final era elaborar um plano comum de tabulações e indicadores, bem como criar uma base de dados única. Com essa finalidade, na primeira etapa do projeto foram criados os seguintes grupos de trabalho: Desenho Conceitual e Temático; Classificações; Capacitação; e Aspectos Tecnológicos, Difusão e Disseminação. Na etapa seguinte, foi criado também o Grupo de Trabalho de Planejamento, Supervisão e Controle.

Considerando-se a heterogeneidade dos formulários censitários utilizados, selecionaram-se os seguintes conjuntos de características para integrar o Censo Comum:

- as referentes às condições de saneamento;
- as demográficas, como idade e sexo; e
- as educacionais, como anos de estudo completos e freqüência escolar.

Para as características econômicas, tinham sido selecionadas a população economicamente ativa e a população ocupada, considerando-se a posição na ocupação e os ramos de atividade, para o que foi necessário criar uma classificação comum de atividades econômicas para o Mercosul. Porém, após o trabalho dos grupos, incorporou-se também a variável ocupação, sendo também definida uma classificação comum de ocupações; Em relação às características migratórias, em particular as migrações entre os países da região, também ocorreu a devida seleção.

2.3.3 Participação do Brasil

A equipe brasileira participou ativamente da coordenação das atividades, desde a concepção técnica do Projeto Censo Comum no Mercosul. O trabalho de geração da base de dados única também foi coordenado pelo Brasil.

Desde o início, o referido projeto recebeu apoio do Fundo de População das Nações Unidas – FNUAP – do Banco Interamericano de Desenvolvimento – BID – do Programa das Nações Unidas para o Desenvolvimento – PNUD – do Instituto Nacional de Estatística e Estudos Econômicos da França – INSEE – do U.S. Census Bureau, do Centro Latinoamericano de Demografia – CELADE – entre outros.

No Brasil, inicialmente o projeto incluía um subprojeto, cujo objetivo era a incorporação de novas tecnologias ao censo brasileiro e, a partir dessa experiência, difundi-las aos outros países do Bloco Mercosul. Nesse sentido, a UNFPA financiou a primeira Prova-Piloto do Censo Demográfico 2000, em dezembro de 1997, onde se ensaiou a captura de dados mediante escâneres e o reconhecimento inteligente de caracteres. Nessa oportunidade, participaram os consultores do Uruguai, que já tinham realizado a primeira experiência de captura de imagens na América Latina, para o censo daquele país, em 1996.

O Brasil ocupou-se em realizar diversas gestões internacionais que permitiram conseguir os financiamentos mencionados acima, além de participar das diversas etapas técnicas do projeto. Participaram dos grupos de trabalho representantes da Diretoria de Informática, do Centro de Documentação e Disseminação de Informações, da Diretoria de Pesquisas e da Coordenação de Acompanhamento e Controle Operacional dos Censos.

2.4 Custos e orçamentos

Os custos de uma operação do porte de um censo demográfico dependem criticamente do grau de planejamento e organização a ela dedicados. As constantes alterações na estrutura de organização do território, decorrentes da instalação de novos municípios, as mudanças nas leis e normas que regulam as áreas de contratação de pessoal e aquisição de bens e serviços, bem como evolução tecnológica, entre outros, são fatores que representam forte impacto no orçamento global de um censo.

Sendo uma operação de longo tempo – abrange vários exercícios fiscais – e de custo muito superior ao das atividades normais do IBGE, é fundamental que os cálculos de seus custos sejam bastante criteriosos e cercados de extremos cuidados, de forma a evitar sub ou superdimensionamentos que venham comprometer sua execução ou impedir/dificultar a aprovação de seu orçamento pelas instâncias superiores competentes.

Já na elaboração das primeiras estimativas de custos, além das referentes aos números de domicílios e pessoas a serem recenseadas – alvos da pesquisa – outros itens determinantes das principais parcelas do custo global da operação precisam estar definidos. Entre esses itens, destacam-se a forma de contratação de pessoal, sua remuneração e encargos e a tecnologia de captura de dados.

Para o Censo Demográfico 2000, o planejamento das grandes ações do governo federal por períodos plurianuais determinou que o IBGE elaborasse estimativas orçamentárias já a partir do ano de 1998. A inclusão do projeto Censo Demográfico 2000 no Plano Plurianual – PPA 2000/2003 – demostrou a vontade e o compromisso do governo federal com a sua realização. Foi também uma forma eficaz de o IBGE garantir junto às instâncias responsáveis pela liberação de recursos, o necessário aporte orçamentário ao longo do período, de forma a não implicar perda de recursos já aplicados, caso houvesse interrupção ou cortes significativos nos valores previamente estimados, sob pena de inviabilizar ou provocar prejuízos irreparáveis aos resultados esperados.

A tabela a seguir informa, ao longo do período abrangido pela operação, a distribuição percentual do orçamento total do Censo Demográfico 2000.

Tabela 2.1 - Distribuição percentual do orçamento, segundo o período da operação - 1998-2003

Período da operação	Distribuição percentual do orçamento (%)
1998	2,79
1999	6,36
2000	80,31
2001	8,32
2002	2,12
2003	0,10

Fonte: Sistema Integrado de Administração Financeira - SIAFI.

Os recursos orçamentários concedidos ao projeto não refletiram, contudo, os gastos totais com a operação, uma vez que a estrutura permanente da instituição, cujos servidores e infra-estrutura foram postos a serviço do censo, se manteve de acordo com seu orçamento normal, sendo, no entanto, de difícil mensuração, visto que as atividades censitárias ocorreram concomitantemente aos trabalhos regulares da instituição. Os gastos da operação foram alocados em seis grandes itens de despesa, como mostra a Tabela 2.2.

Tabela 2.2 - Distribuição percentual do orçamento, segundo o item de despesa

Item de despesa	Distribuição percentual (%)
Pessoal temporário	53,76
Serviços de terceiros	23,96
Investimentos	8,93
Material de consumo e matéria-prima	6,22
Passagens	3,61
Diárias de viagem	3,42
Exercícios anteriores	0,10

Fonte: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, Orçamentos do Censo Demográfico 2000.

Como demonstrado, mais da metade do custo total da operação correspondeu a despesas com pessoal temporário. Por outro lado, mais de 90% dos recursos gastos foram dedicados à operação de coleta dos dados, que, além das despesas com pessoal temporário, envolveu impressão, embalagem e transporte de material, deslocamentos de pessoal, montagem dos postos de coleta, entre outros. A memória de cálculo do orçamento do ano 2000, que, vale lembrar, concentrou 80% do total das despesas da operação, está demonstrada no documento interno *Previsão orçamentária segundo o elemento de despesa para o exercício de 2000* (1999), incluído no CD-ROM que acompanha esta publicação.

Tão importante quanto o aporte suficiente de recursos orçamentários é o cumprimento do cronograma de desembolso, ou seja, o fluxo adequado dos recursos financeiros. Visto que cerca de 60% das despesas do ano da coleta são aplicados no pagamento de pessoal, em especial na remuneração dos Recenseadores, é vital para a garantia da normalidade dos trabalhos de campo que a programação de desembolso ocorra conforme o planejado, de forma a evitar transtornos decorrentes de atrasos de pagamentos, os quais podem trazer prejuízos irreparáveis aos resultados da operação. No Censo Demográfico 2000, cuja coleta de dados teve início no mês de agosto, a programação de desembolso para o ano do recenseamento propriamente dito, o exercício de 2000, foi adequadamente cumprida e ficou assim distribuída:

Tabela 2.3 - Distribuição percentual do desembolso, segundo o mês do exercício de 2000

Mês	Distribuição percentual (%)
Janeiro	1,5
Fevereiro	1,5
Março	2,5
Abril	4,0
Maio	8,0
Junho	10,5
Julho	12,5
Agosto	9,0
Setembro	13,0
Outubro	15,5
Novembro	17,0
Dezembro	5,0

Fonte: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, Orçamentos do Censo Demográfico 2000: exercício 2000.

A Tabela 2.4 demonstra as propostas orçamentárias apresentadas pelo IBGE ao Ministério de Planejamento, Orçamento e Gestão, os valores autorizados pelas leis orçamentárias dos exercícios abrangidos, as dotações suplementares, posteriormente aprovadas pelo Congresso Nacional, as dotações canceladas ou remanejadas, os valores efetivamente executados e seus saldos. Vale destacar que o saldo do exercício de 2000 foi gerado pela dispensa do pagamento do Auxílio-Alimentação aos Recenseadores, por orientação do MPOG, e também pela baixa demanda do Auxílio-Transporte por parte dos contratados mensalistas, cujos valores haviam sido consignados pelo IBGE no orçamento do exercício de 2000.

Tabela 2.4 - Valor das propostas orçamentárias e dotações no período censitário, segundo o tipo de proposta orçamentária e dotação - 1998-2003

Tipo de proposta orçamentária e dotação	Valor (R$)		
	1998	1999	2000
Proposta Orçamentária/orçamento congresso	22 421 653,00	50 854 366,00	390 000 000,00
Dotação inicial aprovado na Lei Orçamentária	19 900 261,00	40 531 650,00	383 951 000,00
Dotação suplementar	-	-	142 713 000,00
Dotação cancelada/remanejada	-	-	-
Total autorizado	19 900 261,00	40 531 650,00	526 664 000,00
Executado	15 796 754,74	36 002 664,71	454 991 349,00
Saldo da dotação	4 103 506,26	4 528 985,29	71 672 651,00

Tipo de proposta orçamentária e dotação	Valor (R$)		
	2001	2002	2003
Proposta Orçamentária/orçamento congresso	51 022 400,00	11 600 000,00	1 215 414,00
Dotação inicial aprovado na Lei Orçamentária	47 103 977,00	11 294 976,00	1 215 414,00
Dotação suplementar	-	-	-
Dotação cancelada/remanejada	1 400 000,00	1 890 997,00	-
Total autorizado	45 703 977,00	9 412 823,00	1 215 414,00
Executado	44 376 761,04	7 855 574,28	1 156 958,62
Saldo da dotação	1 327 215,96	1 557 248,72	58 455,38

Fonte: IBGE, Diretoria Executiva, Coordenação de Planejamento e Supervisão.

2.4.1 Remuneração da coleta

A remuneração dos Recenseadores é outro aspecto de fundamental importância no que diz respeito aos cálculos orçamentários de um censo. Não apenas porque é um dos itens de maior representatividade nos custos da operação, mas também porque a adequada remuneração dessa mão-de-obra é fator importante para assegurar a qualidade da coleta de dados.

Diversas variáveis influenciam na forma e na composição da remuneração dos Recenseadores. Quanto à forma – remuneração fixa mensal ou variável, segundo a produção apresentada – foi adotada no Censo 2000 a remuneração variável, já utilizada em censos anteriores e considerada mais adequada do que a remuneração fixa mensal, porque propicia ao recenseador incentivos de duas ordens:

1ª) fazer a cobertura completa de sua área de trabalho, registrando todas as unidades e pessoas alvo da pesquisa; e

2ª) não prolongar por demais o tempo da coleta, para não prejudicar o valor final de sua remuneração e recebê-la no menor prazo possível, incentivos esses que não ocorreriam, caso a remuneração fosse por retribuição mensal fixa, podendo, pelo contrário, induzir o recenseador a um ritmo de trabalho mais lento, com o intuito de prolongar o tempo do contrato e, por conseqüência, aumentar o período da etapa de coleta e os custos da operação.

2.4.2 Definição das taxas de pagamento dos setores censitários

A opção pelo pagamento da coleta de dados por produção exigiu estudos criteriosos e complexos para definir uma escala de taxas de remuneração de setores censitários e estabelecer os mecanismos para sua aplicação de maneira justa, segundo o grau de dificuldade de realização do trabalho – acesso, percurso, abordagem dos moradores, realização das entrevistas, etc. – em cada região, município e setor censitário, e coerente com os patamares salariais determinados pelos mercados de trabalho locais, de modo que os valores finais pagos para setores censitários urbanos – com quantidades semelhantes de domicílios e pessoas - não fossem iguais, por exemplo, em uma metrópole e em um pequeno município, ou em uma área rural da Região Sul e outra da Região Nordeste.

Para fixar os valores de remuneração da coleta de dados do Censo Demográfico 2000 foi instituída a Tabela Nacional de Taxas para Pagamento da Produção, baixada pela Resolução do Conselho Diretor – R. CD nº 012 A, de 28 de julho de 2000, e a seguir apresentada. Dela constavam 12 alternativas de taxas, numeradas seqüencialmente de 1 a 12, compostas pelos valores unitários de cada item de pagamento, valores esses crescentes, da taxa 1 para a taxa 12, e diferenciados entre os setores urbanos e rurais. Os itens de pagamento estabelecidos foram: unidade visitada, Questionário Básico, pessoa/questionário básico, Questionário da Amostra e pessoa/questionário da amostra.

Tabela 2.5 - Tabela nacional de taxas para pagamento da produção, segundo a situação do domicílio e os itens de pagamento

(continua)

Situação do domicílio e item de pagamento	Taxas (valores em R$)											
	1	2	3	4	5	6	7	8	9	10	11	12
Urbana												
Unidada visitada	0,10	0,11	0,12	0,13	0,15	0,16	0,18	0,19	0,21	0,24	0,26	0,29
Questionário básico	0,28	0,30	0,33	0,37	0,41	0,45	0,49	0,54	0,59	0,65	0,72	0,79
Pessoa/questionário básico	0,10	0,11	0,12	0,14	0,15	0,16	0,18	0,20	0,22	0,24	0,26	0,29
Questionário da amostra	0,81	0,89	0,98	1,08	1,19	1,31	1,44	1,59	1,74	1,92	2,11	2,32
Pessoa/questionário da amostra	0,30	0,33	0,36	0,40	0,44	0,48	0,53	0,58	0,64	0,70	0,78	0,85

Tabela 2.5 - Tabela nacional de taxas para pagamento da produção, segundo a situação do domicílio e os itens de pagamento

(conclusão)

Situação do domicílio e item de pagamento	Taxas (valores em R$)											
	1	2	3	4	5	6	7	8	9	10	11	12
Rural												
Unidada visitada	0,18	0,20	0,22	0,24	0,27	0,29	0,32	0,35	0,39	0,43	0,47	0,52
Questionário básico	0,50	0,55	0,61	0,67	0,74	0,81	0,89	0,98	1,08	1,19	1,30	1,43
Pessoa/questionário básico	0,18	0,19	0,21	0,23	0,26	0,28	0,31	0,34	0,38	0,41	0,45	0,50
Questionário da amostra	1,20	1,32	1,45	1,60	1,76	1,93	2,13	2,34	2,57	2,83	3,11	3,42
Pessoa/questionário da amostra	0,42	0,46	0,51	0,56	0,61	0,67	0,74	0,82	0,90	0,99	1,09	1,19

Fonte: IBGE. Resolução do Conselho Diretor - CD nº 012 A, de 28 de julho de 2000. Concede gratificação por serviços extraordinários para as atividades do Censo Demográfico de 2000. Rio de Janeiro, 2000. Disponível em: <http://w3.aplicacao.de.ibge.gov.br/ Siad/SIAD.nsf>. Acesso em: out. 2003.

Para definir os valores dos itens de pagamento, partiu-se de um valor médio mensal preestabelecido, que foi calculado levando-se em conta a produção média que poderia ser alcançada por um Recenseador que trabalhasse cinco horas por dia e cinco dias por semana a um salário objetivo de R$ 5,00 por hora. Foram utilizadas nesses cálculos, as médias de produção observadas na coleta de Questionários Básico e da Amostra de setores censitários urbanos e rurais das duas Provas-Piloto e do Censo Experimental.

A Tabela Nacional de Taxas para Pagamento da Produção ofereceu 12 diferentes alternativas para pagamento dos setores, que permitiram aos gestores selecionar a taxa mais adequada segundo as características de cada setor, município ou região, observado o limite máximo do recurso orçamentário disponibilizado para tal despesa em cada Unidade da Federação.

Obedecido esse limite orçamentário, era possível também realizar pagamento de determinados setores por **taxa global** – valor fixo predefinido entre o responsável pela atribuição do setor – Supervisor, Agente Censitário Municipal ou o próprio Coordenador de Subárea – e o Recenseador. Esse procedimento se fez necessário para aqueles setores com baixo número de unidades recenseadas, cujo pagamento, pela forma normal, não seria atraente para o Recenseador. Esses setores foram previamente indicados pelas agências do IBGE e tiveram valores máximos definidos, de acordo com as características de cada um, sendo então registrados no sistema de pagamento pelos Chefes das Divisões Estaduais de Pesquisa, a fim de que os pagamentos pudessem ser efetuados em valores iguais ou inferiores àqueles registrados no sistema.

Para o custeio de despesas com locomoção e outras inerentes aos setores de difícil acesso e/ou percurso, como é mais comum nas áreas rurais, foi possível o pagamento de ajuda de custo, cujo valor era fixado pelo Coordenador de Subárea ou por outro servidor do IBGE que recebesse do coordenador tal delegação. A ajuda de custo poderia ser paga antecipadamente, caso isso fosse necessário.

O valor a ser aplicado no pagamento da produção de cada Unidade da Federação foi fixado com base na população estimada, nas características territoriais e nas condições do mercado de trabalho de cada uma delas.

Definida a Tabela Nacional de Taxas para Pagamento da Produção e o valor orçamentário deste item de despesa, foram feitos estudos que sugeriram a taxa mais adequada a cada município, obedecido o teto do recurso orçamentário fixado.

Tal estudo partiu de uma classificação dos municípios por análise discriminante, para a qual foram utilizadas as seguintes variáveis: área do município, área média dos setores, densidade populacional, número médio de domicílios por setor, número de candidato por vaga, densidade populacional do município e percentagem da população rural. Para cada região do País foram previamente estipuladas as taxas que deveriam ser utilizadas, levando-se em conta o salário médio da região.

Assim, foi indicado para cada município que taxa deveria ser aplicada nas áreas urbana e rural. Com base nas estimativas do número de domicílios e de pessoas, foi calculado o valor necessário a cada município para pagamento da coleta e, por conseguinte, o valor necessário a cada Unidade da Federação. Finalmente, considerando-se o valor destacado no Orçamento Geral do Censo para tal despesa, os valores das Unidades da Federação e de seus municípios foram ajustados, substituindo-se algumas taxas determinadas no cálculo inicial.

Após esses estudos iniciais, cada unidade foi informada de seu orçamento para pagamento da coleta e das taxas sugeridas para cada município, as quais deveriam ser previamente analisadas e, se necessário, ajustadas de acordo com outras informações relevantes de conhecimento das instâncias gerenciais locais, mas respeitado o limite orçamentário definido, o qual contemplava uma reserva adicional da ordem de 5%.

Em anexo, apresenta-se a relação das taxas sugeridas e daquelas de fato aplicadas pelas Unidades Regionais no pagamento dos Recenseadores.

2.5 Cronograma geral de atividades

Uma operação do porte e complexidade de um censo demográfico requer controle eficaz, que funcione como a principal ferramenta de gerenciamento de todas as atividades, projetos e subprojetos envolvidos, permitindo o acompanhamento da evolução e do encadeamento das diversas tarefas no tempo e entre as áreas executoras, assim como o gerenciamento de recursos humanos, materiais e financeiros.

A primeira versão do Cronograma Geral de Atividades do Censo Demográfico 2000 foi elaborada ainda no início de 1997, utilizando-se o aplicativo *Microsoft Project*, partindo de uma primeira relação das grandes etapas da operação e suas principais atividades. Com a evolução dos trabalhos, no entanto, surgiu a necessidade de instrumental que operasse em rede, de forma a permitir acesso simultâneo, para consulta e atualização, a todos os gerentes das áreas envolvidas. Como, à época, a versão disponível do *Project* era monousuário, partiu-se para o desenvolvimento de um outro sistema gerencial mais amplo e flexível. Amplo porque almejava abranger diversas informações relevantes sobre cada atividade, projeto e tarefa. Flexível porque deveria permitir a comunicação com outros aplicativos e sistemas, como por exemplo, o Sistema de Indicadores Gerenciais da Coleta, além de gerar relatórios e gráficos, de forma a atender a todas as demandas dos vários níveis gerenciais envolvidos. Levando-se em conta que o IBGE adotara o *software* Lotus Notes como ferramenta de

comunicação, optou-se pelo desenvolvimento de um sistema de gerenciamento de projetos com essa tecnologia, o que, em princípio, deveria facilitar seu uso e a troca de informações entre os gestores do sistema. Contudo, depois de decorrido considerável tempo entre as especificações, o desenvolvimento e inúmeras tentativas de ajustes, a tecnologia foi descartada, em razão da impossibilidade de superação das várias dificuldades surgidas e da falta de funcionalidade do sistema montado.

Retornou-se, então, ao *Microsoft Project*, subdividindo-se a operação em subprojetos independentes, com vistas a facilitar e dar mais agilidade aos processamentos necessários durante os procedimentos de alimentação e atualização dos dados. A subdivisão obedeceu ao critério das grandes atividades, que foram divididas em projetos e suas tarefas, com as interligações, as datas de início e fim, os períodos de tempo demandados – duração - as unidades responsáveis e outras definições internas de critérios e pesos oferecidos pelo *software*.

A COC ficou responsável pela atualização do cronograma, o que era feito em reuniões, de regularidade variada, com os titulares das coordenações e comitês do Censo 2000, conforme as necessidades de acompanhamento ditadas pelas atividades em andamento em cada período da operação.

Apresenta-se em anexo, uma síntese da versão final do Cronograma Geral de Atividades do Censo 2000 – versão resumida da estrutura de projetos e tarefas. A versão completa pode ser visualizada no CD-ROM que acompanha esta publicação.

Referências

AVALIAÇÃO dos censos 96: relatório final. Rio de Janeiro: IBGE, Diretoria de Planejamento e Coordenação, 1998.

BERCOVICH, A. O Censo 2000 no Mercosul: uma experiência transnacional. In: MIGRAÇÕES internacionais: contribuições para políticas. Brasília: Comissão Nacional de População e Desenvolvimento, 2001. p. 275-280. Trabalho apresentado no Seminário sobre Migrações Internacionais - Contribuições Para Políticas, 2000, Brasília.

CAZZOLLI, L. Censo común del Mercosur: ronda del 2000: documento síntesis. In: SEMINÁRIO SOBRE O CENSO 2000 NO MERCOSUL, 8., 1999, Rio de Janeiro. Rio de Janeiro: IBGE, 1999. 20 p.

_____. I Prueba piloto conjunta. In: TALLER DE DISCUSION DEL CENSO DE POBLACION Y VIVIENDAS 2000 EN EL MERCOSUR, 7., 1999, Santiago de Chile. Santiago de Chile: INDEC, 1999.

IBGE. Resolução do Conselho Diretor - CD nº 012 A, de 28 de julho de 2000. Concede gratificação por serviços extraordinários para as atividades do Censo Demográfico de 2000. Rio de Janeiro, 2000. Disponível em: <http://w3.aplicacao.de.ibge.gov.br/Siad/SIAD.nsf>. Acesso em: out. 2003.

PREVISÃO orçamentária segundo o elemento de despesa para o exercício de 2000. Rio de Janeiro: IBGE, Coordenação Operacional dos Censos, 1999.

Anexos

Anexo 2.1 - Percentual de setores pagos, segundo as taxas programadas e as efetivamente utilizadas para pagamento da produção

(continua)

Unidades da Federação	Situação		% dos setores pagos nas taxas											
		Global	1	2	3	4	5	6	7	8	9	10	11	12
Rondônia	Urbana calculado					3,85	84,62	11,54						
	Urbana realizado	5,86				28,47	64,03	1,13	0,10	0,10		0,10	0,10	0,10
	Rural calculado					3,85	48,08	48,08						
	Rural realizado	41,59				0,10		57,63		0,10		0,10	0,10	0,29
Acre	Urbana calculado					31,82	45,45	22,73						
	Urbana realizado					82,26	12,84	4,89						
	Rural calculado					9,09	40,91	50,00						
	Rural realizado	17,09				11,97	24,79	45,73						
Amazonas	Urbana calculado					91,94	6,45	1,61						
	Urbana realizado	1,12				28,05	70,83							
	Rural calculado					3,23	25,81	70,97						
	Rural realizado	32,08				0,18	65,42	2,23	0,09					
Roraima	Urbana calculado						80,00	20,00						
	Urbana realizado						16,12	83,22		0,33			0,33	
	Rural calculado						6,67	93,33						
	Rural realizado	0,50					2,49	97,01						
Pará	Urbana calculado					37,76	51,05	11,19						
	Urbana realizado	0,79				26,82	28,45	42,53	0,28	0,51	0,11	0,11	0,20	0,17
	Rural calculado					18,88	48,95	32,17						
	Rural realizado	13,18	0,08			11,56	60,47	11,15	2,06	1,39		0,04		0,08
Amapá	Urbana calculado					56,25	37,50	6,25						
	Urbana realizado					9,76	23,96	64,20	2,07					
	Rural calculado					6,25	6,25	87,50						
	Rural realizado					8,13	52,85	33,33	5,69					
Tocantins -	Urbana calculado					15,83	76,98	7,19						
	Urbana realizado		2,15	16,91	26,17	18,39	17,72	14,77	0,54	1,61	0,13	1,48	0,13	
	Rural calculado					5,76	49,64	44,60						
	Rural realizado	13,53		0,16	9,18	9,82	18,36	27,05	4,67	4,35	3,86	2,74	0,16	6,12
Maranhão	Urbana calculado					96,77	2,76	0,46						
	Urbana realizado	0,30				68,46	2,85	23,83	0,67	2,37	0,07	1,07		0,22
	Rural calculado					33,18	53,92	12,90						
	Rural realizado	1,32				40,23	38,41	17,03	0,05	1,21	0,37	0,42		0,29

Anexo 2.1 - Percentual de setores pagos, segundo as taxas programadas e as efetivamente utilizadas para pagamento da produção

(continuação)

Unidades da Federação	Situação		% dos setores pagos nas taxas											
		Global	1	2	3	4	5	6	7	8	9	10	11	12
Piauí	Urbana calculado					16,29	74,66	9,05						
	Urbana realizado	0,40				3,87	58,69		37,04					
	Rural calculado					7,69	72,85	19,46						
	Rural realizado	0,91				0,75	70,46	25,26	2,52					
Ceará	Urbana calculado					83,70	13,04	3,26						
	Urbana realizado	0,38				30,07	23,30	5,37	40,86	0,02				
	Rural calculado					24,46	67,39	8,15						
	Rural realizado	0,11				0,22	74,42	25,13	0,04					
Rio Grande do Norte	Urbana calculado					97,59	1,20	1,20						
	Urbana realizado	1,40				56,90	6,84	34,86						
	Rural calculado					26,51	69,28	4,22						
	Rural realizado	12,05				21,50	59,50	6,94						
Paraíba	Urbana calculado					98,21	1,35	0,45						
	Urbana realizado	0,04				37,30	22,73	39,93						
	Rural calculado					17,04	76,68	6,28						
	Rural realizado	1,21				0,06	82,51	16,17						
Pernambuco	Urbana calculado					87,03	10,27	2,70						
	Urbana realizado	2,59				36,12	16,30	2,02	42,95	0,02				
	Rural calculado					20,54	69,19	10,27						
	Rural realizado	6,76				0,06	72,50	16,19	4,46					
Alagoas	Urbana calculado					98,02	0,99	0,99						
	Urbana realizado	0,32				46,01	3,76	40,64	4,99	2,92	0,91	0,26	0,06	0,13
	Rural calculado					39,60	57,43	2,97						
	Rural realizado	1,02				42,70	40,78	5,77	6,05	3,17	0,19	0,19	0,09	
Sergipe	Urbana calculado					30,67	68,00	1,33						
	Urbana realizado	0,07				62,85	37,08							
	Rural calculado					20,00	73,33	6,67						
	Rural realizado	0,11					10,76	89,13						
Bahia	Urbana calculado					94,46	4,82	0,72						
	Urbana realizado	0,76	9,72	0,07	27,76	22,57	10,94	0,13	28,03	0,01				
	Rural calculado					23,86	63,61	12,53						
	Rural realizado	14,61	0,03	27,67	26,32	30,57	0,68	0,09	0,02					

Anexo 2.1 - Percentual de setores pagos, segundo as taxas programadas e as efetivamente utilizadas para pagamento da produção

(continuação)

Unidades da Federação	Situação		% dos setores pagos nas taxas											
		Global	1	2	3	4	5	6	7	8	9	10	11	12
Minas Gerais	Urbana calculado						25,44	71,51	3,05					
	Urbana realizado	0,01	0,02	0,44	0,90	26,39	14,08	17,48	1,96	27,54	1,09	8,29		1,79
	Rural calculado						14,07	73,51	12,43					
	Rural realizado				0,02	0,11	0,13	1,20	4,11	84,31	8,74	0,70	0,28	0,41
Espírito Santo	Urbana calculado						51,95	46,75	1,30					
	Urbana realizado	4,21					94,74	0,61	0,44					
	Rural calculado						42,86	55,84	1,30					
	Rural realizado	10,75					0,10	57,11	31,93	0,10				
Rio de Janeiro	Urbana calculado							46,15	43,96	9,89				
	Urbana realizado	2,44				0,07	2,92	21,21	32,67	37,23	2,10	0,46	0,18	0,70
	Rural calculado							16,48	54,95	28,57				
	Rural realizado	19,61					1,03	14,86	28,48	26,11	1,14	5,57	2,58	0,62
São Paulo	Urbana calculado								32,71	55,35	11,94			
	Urbana realizado	0,53							51,28	30,96	16,70	0,41	0,09	0,02
	Rural calculado								8,06	64,19	27,75			
	Rural realizado	3,20							1,79	63,70	30,14	0,28	0,75	
Paraná	Urbana calculado						15,79	76,19	8,02					
	Urbana realizado	0,20					32,23	34,22	33,34					
	Rural calculado						3,26	70,93	25,81					
	Rural realizado							71,00	28,98					
Santa Catarina	Urbana calculado						2,36	80,89	16,72					
	Urbana realizado	6,16					10,21	55,48	28,15					
	Rural calculado						1,02	73,04	25,94					
	Rural realizado	5,86					0,24	65,27	28,49					
Rio Grande do Sul	Urbana calculado						18,63	73,02	8,35					
	Urbana realizado	0,24		0,31	0,45	3,17	51,13	33,03	9,70	1,31	0,11	0,52	0,01	
	Rural calculado						7,07	68,95	23,98					
	Rural realizado	8,17		0,04		0,02	10,35	52,72	27,75	0,18	0,51	0,20	0,02	0,02
Mato Grosso do Sul	Urbana calculado						2,60	79,22	18,18					
	Urbana realizado	0,27					97,64	1,02	0,54	0,48	0,05			
	Rural calculado						1,30	54,55	44,16					
	Rural realizado	0,12						94,38	0,12	1,99	3,04	0,35		

Anexo 2.1 - Percentual de setores pagos, segundo as taxas programadas e as efetivamente utilizadas para pagamento da produção

(conclusão)

Unidades da Federação	Situação		% dos setores pagos nas taxas											
		Global	1	2	3	4	5	6	7	8	9	10	11	12
Mato Grosso	Urbana calculado						5,56	80,95	13,49					
	Urbana realizado	0,93		0,05		1,66	20,14	38,66	19,11	7,53	3,03	3,27	0,68	4,69
	Rural calculado						0,79	30,16	69,05					
	Rural realizado	24,55		0,15			4,21	21,28	29,41	5,88	2,90	4,43	0,36	4,50
Goiás	Urbana calculado						18,60	75,21	6,20					
	Urbana realizado	0,12	0,02			35,06	15,49	37,18	1,84	5,03	2,59	1,21	0,30	1,12
	Rural calculado						5,79	67,36	26,86					
	Rural realizado	1,84	0,06			0,06	11,70	32,01	21,92	9,41	6,54	6,02	2,93	6,08
Distrito Federal	Urbana calculado										100,00			
	Urbana realizado	3,90								91,92	4,06	0,04	0,04	
	Rural calculado										100,00			
	Rural realizado	62,31									36,92			

Anexo 2.2 - Cronograma geral de atividades do Censo 2000
Versão Resumida

(continua)

Id	Tarefas	Início	Término
1	Estudos iniciais	16/06/97	21/10/99
1.1	Adequação da legislação para contratação de pessoal temporário	02/07/97	24/07/98
1.1.1	Estudo da legislação vigente X implicações decorrentes	02/07/97	19/09/97
1.1.2	Elaboração de proposta de adequação	23/09/97	23/07/98
1.1.3	Envio de proposta ao Ministério da Previdência	24/07/98	24/07/98
1.2	Avaliação das operações censitárias de 1996	02/07/97	17/05/99
1.2.1	Elaboração de relatórios estaduais	02/07/97	29/01/98
1.2.2	Reuniões nas URs	20/02/98	30/05/98
1.2.3	Organização de relatórios e envio para as áreas técnicas	01/06/98	26/06/98
1.2.4	Avaliação e síntese	09/07/98	10/07/98
1.2.5	Consolidação e relatório final	14/05/99	17/05/99
1.3	Projetos para financiamentos externos	27/05/98	30/09/98
1.3.1	Implementação de novas tecnologias na Captura de Dados (1ª Prova Piloto)	27/05/98	29/06/98
1.3.2	II Seminário sobre o Censo 2000 no Mercosul	27/07/98	06/08/98
1.3.3	Incorporação de novas metodologias e integração com o Mercosul	01/09/98	30/09/98
1.4	Definição do programa de trabalho	16/06/97	10/11/98
1.4.1	Definição preliminar do conteúdo dos questionários	28/11/97	26/05/98
1.4.2	Definição das equipes de trabalho	30/06/98	24/07/98
1.4.3	Outras definições preliminares	16/06/97	26/11/97
1.4.4	Amostragem	27/05/98	10/11/98
1.5	Consulta aos usuários	16/06/97	22/04/99
1.5.1	Desenho inicial do formulário	16/06/97	08/08/97
1.5.2	Experiências internacionais	28/08/98	09/10/98
1.5.3	Solicitações do Mercosul	28/08/98	25/09/98
1.5.4	Necessidades do Governo	11/08/97	03/10/97
1.5.5	Consultas aos usuários internos	12/10/98	03/11/98
1.5.6	Reunião com a Comissão Consultiva	03/11/98	04/11/98
1.5.7	Reuniões temáticas com usuários	05/11/98	06/11/98
1.5.8	Incorporação das modificações	05/11/98	10/11/98
1.5.9	Consultas aos usuários externos	11/11/98	22/04/99
1.6	Adequação da legislação da Divisão Territorial Brasileira	01/10/98	21/10/99
1.6.1	Levantamento das inconsistências	01/10/98	09/06/99
1.6.2	Elaboração de proposta para tratamento da DTB para fins de censos	10/06/99	15/09/99
1.6.3	Encaminhamento de proposta ao TCU	18/10/99	21/10/99
1.7	Definição de tecnologia e sistemas de entrada de dados	15/09/97	22/04/98
2	Reuniões da Comissão de Planejamento e Organização do Censo 2000	29/07/97	15/12/03
2.1	Instalação	29/07/97	29/07/97
2.2	Primeira reunião da CPO 2000	08/08/97	11/08/97
2.3	Segunda reunião da CPO 2000	15/08/97	18/08/97
2.4	Terceira reunião da CPO 2000	04/09/97	05/09/97
2.5	Quarta reunião da CPO 2000	19/09/97	22/09/97
2.6	Quinta reunião da CPO 2000	08/10/97	09/10/97
2.7	Sexta reunião da CPO 2000	06/11/97	07/11/97
2.8	Sétima reunião da CPO 2000	25/11/97	26/11/97
2.9	Oitava reunião da CPO 2000(primeira parte)	02/03/98	03/03/98
2.10	Oitava reunião da CPO 2000(segunda parte)	09/03/98	10/03/98
2.11	Nona reunião da CPO 2000	14/04/98	15/04/98
2.12	Décima reunião da CPO 2000	21/04/98	22/04/98
2.13	Décima primeira reunião da CPO 2000	19/05/98	20/05/98
2.14	Décima segunda reunião da CPO 2000	21/05/98	22/05/98
2.15	Décima terceira reunião CPO 2000	02/06/98	03/06/98
2.16	Décima quarta reunião da CPO 2000	04/08/98	05/08/98
2.17	Décima quinta reunião da CPO 2000	19/08/98	20/08/98
2.18	Décima sexta reunião da CPO 2000	07/09/98	08/09/98
2.19	Décima sétima reunião da CPO 2000	28/09/98	29/09/98

Anexo 2.2 - Cronograma geral de atividades do Censo 2000
Versão Resumida

(continuação)

Id	Tarefas	Início	Término
2.20	Décima oitava reunião da CPO 2000	26/11/98	27/11/98
2.21	Décima nona reunião da CPO 2000	10/12/98	11/12/98
2.22	Vigésima reunião da CPO 2000	18/12/98	21/12/98
2.23	Vigésima segunda reunião da CPO 2000	25/02/99	26/02/99
2.24	Vigésima primeira reunião da CPO 2000	04/01/99	05/01/99
2.25	Vigésima terceira reunião da CPO 2000	14/04/99	15/04/99
2.26	Vigésima quarta reunião da CPO 2000	23/04/99	26/04/99
2.27	Vigésima quinta reunião da CPO 2000	13/05/99	14/05/99
2.28	Vigésima sexta reunião da CPO 2000	17/05/99	18/05/99
2.29	Vigésima sétima reunião da CPO 2000	13/08/99	16/08/99
2.30	Vigésima oitava reunião da CPO 2000	15/09/99	16/09/99
2.31	Vigésima nona reunião da CPO 2000	07/10/99	07/10/99
2.32	Encontro Regional de Belém	28/10/99	28/10/99
2.33	Trigésima reunião da CPO 2000 - Encontro Regional de Belém	09/11/99	09/11/99
2.34	Trigésima primeira reunião da CPO 2000	19/11/99	22/11/99
2.35	Trigésima segunda reunião da CPO 2000	03/12/99	06/12/99
2.36	Trigésima terceira reunião da CPO 2000	06/01/00	07/01/00
2.37	Encontro Regional de Fortaleza	21/01/00	21/01/00
2.38	Encontro Regional de Cuiabá	04/02/00	04/02/00
2.39	Trigésima quarta reunião da CPO 2000 - Encontro Regional de Cuiabá	16/03/00	17/03/00
2.40	Encontro Regional de Curitiba	24/03/00	24/03/00
2.41	Trigésima quinta reunião da CPO 2000	30/03/00	31/03/00
2.42	Trigésima sexta reunião da CPO 2000	06/04/00	07/04/00
2.43	Trigésima sétima reunião da CPO 2000	19/04/00	19/04/00
2.44	Trigésima oitava reunião da CPO 2000 - Encontro regional de Curitiba	05/05/00	05/05/00
2.45	Encontro com as Unidades Regionais em Conservatória	12/05/00	12/05/00
2.46	Trigésima nona reunião da CPO 2000	29/05/00	29/05/00
2.47	Quadragésima reunião da CPO 2000	20/06/00	20/06/00
2.48	Quadragésima primeira reunião da CPO 2000	29/06/00	30/06/00
2.49	Quadragésima segunda reunião da CPO 2000	14/07/00	17/07/00
2.50	Quadragésima terceira reunião da CPO 2000	27/07/00	28/07/00
2.51	Quadragésima quarta reunião da CPO 2000	11/08/00	14/08/00
2.52	Quadragésima quinta reunião da CPO 2000	25/08/00	28/08/00
2.53	Quadragésima sexta reunião da CPO 2000	08/09/00	11/09/00
2.54	Quadragésima sétima reunião da CPO 2000	24/09/00	25/09/00
2.55	Quadragésima oitava reunião da CPO 2000	06/10/00	09/10/00
2.56	Quadragésima nona reunião da CPO 2000	20/10/00	23/10/00
2.57	Quinquagésima reunião da CPO 2000	03/11/00	06/11/00
2.58	Quinquagésima primeira reunião da CPO 2000	20/10/00	23/10/00
2.59	Quinquagésima segunda reunião da CPO 2000	01/12/00	04/12/00
2.60	Quinquagésima terceira reunião da CPO 2000	15/12/00	18/12/00
2.61	Quinquagésima quarta reunião da CPO 2000	03/04/00	03/04/00
2.62	Quinquagésima quinta reunião da CPO 2000	03/04/00	03/04/00
2.63	Quinquagésima sexta reunião da CPO 2000	03/04/00	03/04/00
2.64	Quinquagésima sétima reunião da CPO 2000	03/04/00	03/04/00
2.65	Quinquagésima oitava reunião da CPO 2000	03/04/00	03/04/00
3	Definição dos conceitos e testes para a operação censitária	08/08/97	27/12/00
3.1	Primeira Prova Piloto	08/08/97	16/03/99
3.1.1	Conteúdo do questionário	08/08/97	08/08/97
3.1.2	Seleção das UF / Municípios participantes	27/08/97	30/09/97
3.1.3	Questionários	20/10/97	20/10/97
3.1.4	Manual do recenseador	24/09/97	24/09/97
3.1.5	Relatório do recenseador	18/08/97	18/08/97
3.1.6	Operação de campo	24/11/97	12/12/97

Anexo 2.2 - Cronograma geral de atividades do Censo 2000
Versão Resumida

(continuação)

Id	Tarefas	Início	Término
3.1.7	Acompanhamento da coleta	03/12/97	15/12/97
3.1.8	Processamento	15/12/97	11/03/98
3.1.9	Avaliação da operação de campo	02/02/98	16/03/98
3.1.10	Análise dos relatórios de avaliação	02/02/99	16/03/99
3.1.11	Avaliação dos resultados	17/03/98	17/04/98
3.2	Segunda Prova Piloto	16/02/98	18/03/99
3.2.1	Planejamento	04/03/98	19/08/98
3.2.2	Conteúdo dos questionários	16/02/98	22/06/98
3.2.3	Instrumentos de coleta	24/06/98	27/08/98
3.2.4	Manuais e relatórios	04/05/98	27/08/98
3.2.5	Treinamento	01/06/98	08/10/98
3.2.6	Operação de campo	10/08/98	29/09/98
3.2.7	Prova conjunta com o Mercosul	13/07/98	18/03/99
3.3	Censo Experimental	01/04/98	16/08/00
3.3.1	Sistema de Administração de Pessoal	23/10/98	30/07/99
3.3.2	Sistema Gerencial de Controle de Suprimento – Fluxo de Recursos	01/01/99	03/04/00
3.3.3	Recrutamento e seleção	01/03/99	30/11/99
3.3.4	Definição das frações amostrais	26/04/99	26/04/99
3.3.5	Conteúdo dos questionários	01/09/98	10/05/99
3.3.6	Instrumentos de coleta	29/10/98	11/08/99
3.3.7	Operação de campo	02/08/99	28/10/99
3.3.8	Definição do sistema de empastamento e dos modelos das pastas	23/04/99	20/05/99
3.3.9	Processamento	05/05/98	16/08/00
3.3.10	Pesquisa de Avaliação	01/04/98	25/11/99
3.3.11	Coleta de Dados	01/09/99	11/04/00
3.4	Avaliação e ajustes pós-Censo Experimental	26/07/99	27/12/00
3.4.1	Instrumentos de coleta	01/11/99	01/12/99
3.4.2	Manuais	01/11/99	15/12/99
3.4.3	Treinamento	01/11/99	23/12/99
3.4.4	Sistema de Captura de Dados	03/07/00	09/11/00
3.4.5	Sistema de Administração de Pessoal	26/07/99	28/01/00
3.4.6	Sistema de Indicadores Gerenciais da Coleta	28/09/99	30/06/00
3.4.7	Sistema de Codificação Assistida	03/07/00	28/09/00
3.4.8	Sistema de Imputação de Dados	03/07/00	28/09/00
3.4.9	Sistema de Expansão dos Dados da Amostra	01/08/00	28/09/00
3.4.10	Sistema de Tabulação dos Dados	01/09/00	27/12/00
4	Divulgação do Censo 2000	20/08/97	27/12/02
4.1	Ações promocionais de divulgação	20/08/97	27/12/02
4.1.1	Definição da identidade visual – logomarca	20/08/97	15/10/98
4.1.2	Internet	09/08/99	27/12/02
4.1.3	Desenvolvimento da página do Censo 2000	03/01/00	01/03/00
4.1.4	Intranet	25/05/00	06/06/00
4.1.5	Revista do Censo 2000 (Vou Te Contar)	01/07/99	14/03/02
4.2	Diagramação e impressão de material promocional	02/09/97	01/01/01
4.2.1	Impressão das etiquetas	08/11/99	31/01/00
4.2.2	Folheto	02/09/97	25/07/00
4.2.3	Cartaz promocional	02/09/97	24/07/00
4.2.4	Distribuição do material promocional	02/01/98	01/01/01
4.3	Comissões Censitárias Municipais	03/05/99	29/12/00
4.3.1	Desenvolvimento do projeto	03/05/99	10/12/99
4.3.2	Implantação	10/01/00	28/04/00
4.3.3	Realização das reuniões	01/02/00	30/11/00
4.3.4	Consolidação dos relatórios	30/10/00	13/11/00
4.3.5	Relatório final	14/11/00	29/12/00

Anexo 2.2 - Cronograma geral de atividades do Censo 2000
Versão Resumida

(continuação)

Id	Tarefas	Início	Término
4.4	Projeto Vamos Contar	22/11/99	29/11/00
4.4.1	VC 01 Mapa Demográfico 1 / Político - frente e verso	01/12/99	19/07/00
4.4.2	VC 02 Mapa Demográfico 2 / Político - frente e verso	01/12/99	19/07/00
4.4.3	VC 03 Carta ao Professor	01/12/99	19/07/00
4.4.4	VC 04 Carta ao Diretor	01/12/99	19/07/00
4.4.5	VC 05 Carta ao Conselho Escola - Comunidade	01/12/99	19/07/00
4.4.6	VC 06 Cartaz	01/12/99	19/07/00
4.4.7	VC 07 Folder	01/12/99	19/07/00
4.4.8	VC 08 Guia do Professor - Ensino Fundamental 1 a 4 série	01/12/99	19/07/00
4.4.9	VC 09 Guia do Professor - Ensino Fundamental 5 a 8 série	01/12/99	19/07/00
4.4.10	VC 10 Guia do Professor - Ensino Médio	01/12/99	19/07/00
4.4.11	VC 11 Ficha de Avaliação	01/12/99	19/07/00
4.4.12	VC 12 Produção do vídeo - instruções de conteúdo	01/06/00	19/07/00
4.4.13	VC 13 Produção para programação na televisão	22/11/99	19/07/00
4.4.14	VC 14 Página na Internet	29/11/99	19/07/00
4.4.15	VC 15 Caixa 1	01/03/00	19/07/00
4.4.16	VC 16 Caixa 2	01/03/00	19/07/00
4.4.17	VC 17 Canudo / Mapa	01/03/00	19/07/00
4.4.18	Material de acondicionamento	01/12/99	10/07/00
4.4.19	Distribuição para as escolas	25/09/00	29/11/00
4.5	Integração com as comunidades (acadêmica, étnica ou religiosa)	03/01/00	26/12/00
4.6	Divulgação em instituições públicas e privadas	31/03/00	30/01/01
4.7	Publicidade e propaganda	09/08/99	05/11/01
4.7.1	Elaboração do briefing	09/08/99	29/09/99
4.7.2	Preparo do edital de licitação	04/10/99	25/10/99
4.7.3	Parecer da PGE	26/10/99	09/11/99
4.7.4	Licitação e contratação	30/11/99	04/04/00
4.7.5	Criação	05/04/00	30/05/00
4.7.6	Veiculação antes do início da coleta	29/07/00	31/07/00
4.7.7	Veiculação após o início da coleta	01/08/00	05/11/01
5	Base Operacional e Geográfica – BOG	02/01/97	30/12/03
5.1	Trabalhos Iniciais	02/01/97	30/11/99
5.1.1	Estudo e definição do projeto para revisão da BOG	02/01/97	10/12/97
5.1.2	Reuniões técnicas com as URs para avaliação e integração	15/04/97	30/11/99
5.1.3	Contratação de pessoal para o preparo da BOG	02/07/97	07/08/98
5.2	Aquisições e preparo da infra-estrutura	27/02/98	19/07/01
5.2.1	Aquisição de microcomputadores para a Diretoria de Geociências	27/09/99	15/02/00
5.2.2	Aquisição de material de informática para a Diretoria de Geociências	25/10/99	09/06/00
5.2.3	Locação de "plotters"	06/04/00	19/07/01
5.2.4	Aquisição de microcomputadores p/ setores de Base Operacional	26/03/99	15/06/99
5.2.5	Obras e instalações	27/02/98	03/07/98
5.2.6	Contratação de serviços	23/03/98	08/10/99
5.3	Atualização das áreas urbanas dos municípios brasileiros	02/07/97	28/12/00
5.3.1	Capacitação e treinamento de equipes das URs (servidores do IBGE)	27/01/98	23/06/00
5.3.2	Repasse de instruções de serviço	03/04/98	30/06/00
5.3.3	Acompanhamento e supervisão das atividades da vertente urbana	02/06/98	28/12/00
5.3.4	Inventário dos documentos existentes	02/07/97	28/05/99
5.3.5	Aquisição de documentos	02/07/97	30/06/99
5.3.6	Avaliação do grau de desatualização	11/08/97	30/07/99
5.3.7	Atividades de gabinete	11/08/97	30/08/99
5.3.8	Atualização de campo (organização e execução das operações)	22/04/98	30/09/99
5.3.9	Elaboração do Mapa de Localidade – ML	04/05/98	30/12/99
5.3.10	Planejamento da Malha Setorial Urbana 2000	04/05/98	30/09/99
5.3.11	Elaboração do Mapa de Localidade Estatístico – MLE	02/06/98	28/04/00

Anexo 2.2 - Cronograma geral de atividades do Censo 2000
Versão Resumida

(continuação)

Id	Tarefas	Início	Término
5.3.12	Revisão do Planejamento da Malha Setorial Urbana 2000	01/10/99	31/03/00
5.4	Atualização das áreas rurais dos municípios brasileiros – vertente rural	10/06/97	29/12/00
5.4.1	Capacitação e treinamento de equipes das URs (servidores do IBGE)	10/06/97	16/07/99
5.4.2	Repasse de instruções de serviço	27/01/98	30/03/00
5.4.3	Acompanhamento e supervisão das atividades da vertente rural	02/03/98	29/12/00
5.4.4	Inventário dos documentos existentes	02/07/97	28/05/99
5.4.5	Aquisição de documentos	02/07/97	30/06/99
5.4.6	Avaliação do grau de desatualização	04/08/97	04/08/97
5.4.7	Atividades de gabinete	04/08/97	30/08/99
5.4.8	Atualização de campo (organização e execução das operações)	22/04/98	30/09/99
5.4.9	Cadastro das Fichas de Atualização de Campo	02/03/98	30/12/99
5.4.10	Elaboração do Mapa Municipal (MM)	01/04/98	29/12/00
5.4.11	Planejamento da Malha Setorial Rural 2000	04/05/98	30/09/99
5.4.12	Elaboração do Mapa Municipal Estatístico (MME)	04/06/98	28/12/00
5.4.13	Revisão do planejamento da Malha Setorial Rural 2000	01/10/99	30/03/00
5.5	Avaliação de novas unidades municipais	13/05/98	28/04/00
5.6	Sistema Alfanumérico	02/06/97	29/03/01
5.6.1	Cadastro de Segmentos de Logradouros	04/09/97	29/03/01
5.6.2	Espelho de Setores (GR-7.04) e Descrição de Setores (GR-7.03), Tabelas DTB	02/06/97	29/09/98
5.6.3	Divisão Político-Administrativa (BET- Banco de Estruturas Territoriais)	01/03/99	31/07/00
5.7	Áreas Geográficas para Disseminação	08/09/97	30/12/03
5.7.1	Estudo metodológico	08/09/97	31/07/98
5.7.2	Desenvolvimento de sistemas	16/02/98	29/07/99
5.7.3	Teste e validação dos sistemas	02/08/99	26/08/99
5.7.4	Exploração e análise de dados estatísticos	27/02/98	28/12/98
5.7.5	Teste para validação de conceitos e critérios para delimitação de áreas	05/01/99	12/01/00
5.7.6	Delimitação final das áreas	10/01/01	31/12/01
5.7.7	Caracterização dos Setores Censitários de Baixa Renda	01/10/98	30/12/03
5.7.8	Divisão Meso e Micro Regional do Brasil	04/09/00	28/02/02
5.8	Preparo dos instrumentos de coleta	24/09/99	30/06/00
5.8.1	Preparo dos mapas de setores	24/09/99	30/06/00
5.8.2	Impressão das descrições de setores	07/10/99	31/03/00
5.8.3	Montagem da Caderneta do Setor	03/04/00	30/06/00
5.8.4	Impressão dos relatórios do cadastro de segmentos de logradouros	17/04/00	22/05/00
5.9	Revisão final	01/06/00	28/07/00
5.9.1	Avaliação e validação dos instrumentos de coleta	01/06/00	30/06/00
5.9.2	Correções / atualizações finais	01/06/00	03/07/00
5.9.3	Reemissão dos instrumentos de coleta para os setores com problemas	15/06/00	28/07/00
5.10	Monitoramento da coleta	01/08/00	29/12/00
5.10.1	Acompanhamento do Arquivo Alfanumérico	01/08/00	28/12/00
5.10.2	Acompanhamento da Divisão Político-administrativa	01/08/00	28/12/00
5.10.3	Compatibilização da malha municipal (vertente rural)	01/08/00	29/12/00
5.10.4	Avaliação dos trabalhos de atualização da Base Territorial (vertente rural)	01/09/00	03/10/00
6	Recrutamento, seleção e contratação de pessoal	10/05/99	23/11/00
6.1	Elaboração do projeto	10/05/99	06/09/99
6.1.1	Apresentação à CPO	10/05/99	06/09/99
6.2	Processo seletivo para Analista Censitário da administração central	24/05/99	22/10/99
6.2.1	Solicitação de autorização	24/05/99	25/05/99
6.2.2	Autorização	25/05/99	09/07/99
6.2.3	Publicação do Edital	12/07/99	16/07/99
6.2.4	Inscrições	02/08/99	13/08/99
6.2.5	Confirmação de inscrição	20/08/99	26/08/99
6.2.6	Aplicação da prova	10/09/99	12/09/99
6.2.7	Divulgação dos gabaritos	13/09/99	14/09/99

Anexo 2.2 - Cronograma geral de atividades do Censo 2000
Versão Resumida

(continuação)

Id	Tarefas	Início	Término
6.4.2	Exame do Edital pela Procuradoria Geral	06/10/99	13/10/99
6.4.3	Publicação do Edital	04/11/99	05/11/99
6.4.4	Preparo das propostas	08/11/99	06/12/99
6.4.5	Apresentação das propostas e habilitação	06/12/99	07/12/99
6.4.6	Exames dos preços	07/12/99	08/12/99
6.4.7	Julgamento	08/12/99	09/12/99
6.4.8	Prazo para recursos	12/11/99	15/11/99
6.4.9	Contratação da empresa	16/12/99	03/01/00
6.4.10	Contratação da ECT	06/10/99	31/01/00
6.5	Processo seletivo para a contratação de ACM e Supervisor	06/10/99	08/06/00
6.5.1	Procedimentos Iniciais	06/10/99	20/01/00
6.5.2	Cadastramento dos Locais das Prova	16/11/99	17/02/00
6.5.3	Inscrição de Candidatos	29/11/99	14/03/00
6.5.4	Alocação dos Candidatos	15/03/00	17/03/00
6.5.5	Cartões de Confirmação de Inscrição	17/03/00	05/04/00
6.5.6	Divulgação dos Locais de Prova	24/03/00	11/04/00
6.5.7	Preparo e Distribuição da Prova e do Material Administrativo	12/01/00	31/03/00
6.5.8	Aplicação da Prova	20/03/00	20/04/00
6.5.9	Correção das Provas e Emissão das Listagens de Resultados	24/04/00	02/05/00
6.5.10	Divulgação dos Resultados	02/05/00	18/05/00
6.5.11	Contratação	15/05/00	08/06/00
6.6	Processo seletivo para Recenseador	01/02/00	31/07/00
6.6.1	Procedimentos iniciais	01/02/00	17/03/00
6.6.2	Cadastramento dos locais de prova	01/02/00	07/04/00
6.6.3	Inscrição de candidatos	04/02/00	26/04/00
6.6.4	Alocação dos candidatos	26/04/00	28/05/00
6.6.5	Cartões de confirmação de inscrição	02/05/00	24/05/00
6.6.6	Preparo e distribuição da prova e do material administrativo	25/04/00	19/05/00
6.6.7	Aplicação da prova	10/04/00	06/06/00
6.6.8	Correção da prova e emissão das listagens de resultados	29/05/00	23/06/00
6.6.9	Divulgação dos resultados	23/06/00	31/07/00
6.6.10	Contratação	28/07/00	31/07/00
6.7	Processo seletivo para pessoal dos Centros de Captura de Dados	02/05/00	23/11/00
6.7.1	Elaboração do Projeto	02/05/00	15/05/00
6.7.2	Validação do Projeto com as Chefias dos Centros	16/05/00	22/05/00
6.7.3	Seleção da Empresa Executora	17/05/00	25/05/00
6.7.4	Elaboração do Edital	01/08/00	07/08/00
6.7.5	Parecer da Procuradoria Geral sobre o Edital	08/08/00	11/08/00
6.7.6	Publicação do Edital	14/08/00	16/08/00
6.7.7	Elaboração do Treinamento para Seleção de Supervisores	01/08/00	13/10/00
6.7.8	Divulgação do Processo Seletivo Simplificado	14/08/00	06/09/00
6.7.9	Inscrição	28/08/00	06/09/00
6.7.10	Aplicação da Prova	24/09/00	24/09/00
6.7.11	Divulgação de Resultados	09/10/00	11/10/00
6.7.12	Seleção de Supervisores	16/10/00	03/11/00
6.7.13	Convocação para Contrato	06/11/00	23/11/00
7	Treinamento para a coleta	25/05/98	13/11/00
7.1	Desenvolvimento do projeto e preparo do material	25/05/98	25/10/00
7.1.1	Formação do Grupo de Trabalho	25/05/98	05/06/98
7.1.2	Reuniões p/ planejamento, elaboração e acompanhamento do projeto	19/06/98	25/10/00
7.1.3	Contatos com especialistas de empresas	19/06/98	11/09/98
7.1.4	Definição do modelo e elaboração do projeto piloto de treinamento	03/08/98	26/02/99
7.1.5	Contratação de serviços de terceiros para desenvolvimento	01/04/99	31/01/00
7.1.6	Elaboração do material instrucional e normas de operacionalização	01/02/99	18/05/99

Anexo 2.2 - Cronograma geral de atividades do Censo 2000
Versão Resumida

(continuação)

Id	Tarefas	Início	Término
7.1.7	Revisão do material instrucional e normas de operacionalização	15/10/99	20/12/99
7.1.8	Produção de vídeos	15/04/99	24/03/00
7.1.9	Avaliação do projeto pós-Censo Experimental	02/08/99	30/09/99
7.1.10	Ajustes e revisão nos instrumentos	15/10/99	30/12/99
7.2	Editoração e diagramação do material	09/07/99	09/06/00
7.2.1	AD 01 Manual do Sistema de Administração de Pessoal - Orientações Gerais	17/01/00	14/03/00
7.2.2	CD 1.01Questionário Básico	08/10/99	11/01/00
7.2.3	CD 1.02 Questionário da Amostra	08/10/99	11/01/00
7.2.4	CD 1.03 Folha de Domicílio Coletivo	08/10/99	18/02/00
7.2.5	CD 1.06 Caderneta do Setor	08/10/99	09/03/00
7.2.6	CD 1.07 Folha de Coleta	08/10/99	15/05/00
7.2.7	CD 1.09 Manual do Recenseador	30/12/99	01/03/00
7.2.8	CD 1.10 Manual do Supervisor	30/12/99	28/02/00
7.2.9	CD 1.11 Manual do Agente Censitário Municipal	04/01/00	28/02/00
7.2.10	CD 1.12 Manual de Coordenação Técnica	04/01/00	01/03/00
7.2.11	CD 1.13 Guia de remessa de questionários	22/12/99	03/01/00
7.2.12	CD 1.15 Agenda do Supervisor	22/12/99	07/02/00
7.2.13	TR 01 Manual do Instrutor	21/01/00	14/03/00
7.2.14	TR 02 Roteiro de Estudo para o Recenseador	17/11/99	13/03/00
7.2.15	TR 03 Teste Inicial	17/01/00	02/03/00
7.2.16	TR 04 Caderno de Exercícios	17/01/00	17/03/00
7.2.17	TR 05 Teste Final G1	30/11/99	03/01/00
7.2.18	TR 06 Teste Final A	17/01/00	20/03/00
7.2.19	TR 07 Teste Final B	17/01/00	20/03/00
7.2.20	TR 08 Gabarito do Teste Final G1	10/12/99	03/01/00
7.2.21	TR 09 Gabarito do Teste Final A	17/01/00	20/03/00
7.2.22	TR 10 Gabarito do Teste Final B	17/01/00	20/03/00
7.2.23	TR 11 Avaliação do Treinamento - Treinando	31/01/00	10/03/00
7.2.24	TR 12 Avaliação do Treinamento - Instrutor	31/01/00	10/03/00
7.2.25	TR 13 Avaliação do Treinamento - Observador	31/01/00	10/03/00
7.2.26	TR 14 Álbum Seriado	27/12/99	14/02/00
7.2.27	TR 17 Roteiro de Aulas para Treinamento de 12 horas	31/01/00	09/06/00
7.2.28	TR 18 Manual do Instrutor - Base Operacioanl Geográfica	01/03/00	21/03/00
7.2.29	TR 19 - Caderno de Exercícios - Base Operacional Geográfica	17/01/00	28/02/00
7.2.30	TR 20 Álbum Seriado II	21/01/00	14/03/00
7.2.31	TR 21 Avaliação do Treinamento - Treinando	31/01/00	10/03/00
7.2.32	TR 22 Avaliação do Treinamento - Instrutor	31/01/00	10/03/00
7.2.33	TR 23 - Instruções para Revisão do Trabalho da Base Operacional Geográfica	14/01/00	28/03/00
7.2.34	TR 24 Manual do Usuário - Sistema de Indicadores Gerencias da Coleta de Dados para o Censo 2000	17/01/00	28/03/00
7.2.35	TR 25 Fitas de vídeo Censo 2000 Treinamento	20/03/00	14/04/00
7.2.36	TR 26 Fitas de vídeo Censo 2000 Institucional	01/02/00	17/02/00
7.2.37	TR 27 Pasta para treinamento	09/07/99	15/07/99
7.2.38	TR 31 Avaliação de treinamento - observador	31/01/00	10/03/00
7.3	Impressão do material	03/12/99	07/07/00
7.3.1	AD 01 Manual do Sistema de Administração de Pessoal - Orientações Gerais	20/03/00	03/05/00
7.3.2	CD 1.01Questionário Básico	03/12/99	11/01/00
7.3.3	CD 1.02 Questionário da Amostra	03/12/99	07/07/00
7.3.4	CD 1.03 Folha de Domicílio Coletivo	03/12/99	11/01/00
7.3.5	CD 1.06 Caderneta do Setor	03/12/99	11/01/00
7.3.6	CD 1.07 Folha de Coleta	03/12/99	11/01/00
7.3.7	CD 1.09 - Manual do Recenseador	03/03/00	07/07/00
7.3.8	CD 1.10 - Manual do Supervisor	06/03/00	30/03/00
7.3.9	CD 1.11 - Manual do ACM	06/03/00	30/03/00

Anexo 2.2 - Cronograma geral de atividades do Censo 2000
Versão Resumida

(continuação)

Id	Tarefas	Início	Término
7.3.10	CD 1.12 - Manual de Coordenação	06/03/00	30/03/00
7.3.11	CD 1.13 - Guia de Remessa de Questionários	03/01/00	31/01/00
7.3.12	CD 1.15 - Agenda do Supervisor	14/02/00	13/03/00
7.3.13	TR 01 Manual do Instrutor	15/03/00	29/06/00
7.3.14	TR 02 Roteiro de Estudo para o Recenseador	14/03/00	07/07/00
7.3.15	TR 03 Teste Inicial	14/03/00	07/07/00
7.3.16	TR 04 Caderno de Exercícios	20/03/00	07/07/00
7.3.17	TR 05 Teste Final G1	03/01/00	31/01/00
7.3.18	TR 06 Teste Final A	22/03/00	26/06/00
7.3.19	TR 07 Teste Final B	22/03/00	05/07/00
7.3.20	TR 08 Gabarito do Teste Final G1	03/01/00	31/01/00
7.3.21	TR 09 Gabarito do Teste Final A	22/03/00	30/03/00
7.3.22	TR 10 Gabarito do Teste Final B	04/01/00	30/03/00
7.3.23	TR 11 Avaliação do Treinamento - Treinando	16/03/00	30/03/00
7.3.24	TR 12 Avaliação do Treinamento - instrutor	16/03/00	30/03/00
7.3.25	TR 13 Avaliação do Treinamento - observador	16/03/00	30/03/00
7.3.26	TR 14 Álbum Seriado	16/02/00	07/07/00
7.3.27	TR 15 A/B Instrumentos Ampliados - Folha de Coleta	16/03/00	30/03/00
7.3.28	TR 16 Instrumentos Ampliados - Folha de Domicílio Coletivo	16/03/00	30/03/00
7.3.29	TR 17 Roteiro das Aulas para Treinamentos de 12 horas	26/06/00	03/07/00
7.3.30	TR 18 Manual do Instrutor	13/03/00	30/03/00
7.3.31	TR 19 Caderno de Exercícios - Base Operacional Geográfica	01/03/00	30/03/00
7.3.32	TR 20 Álbum Seriado II	16/03/00	30/03/00
7.3.33	TR 21 Avaliação de treinamento - Treinando	14/03/00	30/03/00
7.3.34	TR 22 Avaliação do Treinamento - Instrutor	15/03/00	30/03/00
7.3.35	TR 23 Instruções para Revisão do Trabalho da Base Operacional Geográfica	28/03/00	28/04/00
7.3.36	TR 24 Manual do Usuário - Sistema de Indicadores Gerenciais da Coleta de Dados	28/03/00	28/04/00
7.3.37	TR 25 Vídeo Censo 2000 Treinamento	20/03/00	30/03/00
7.3.38	TR 26 Vídeo Censo 2000 Institucional	10/02/00	10/03/00
7.3.39	TR 27 - Pasta para treinamento	14/02/00	08/06/00
7.3.40	TR 31 Avaliação do Treinamento - Observador	15/03/00	30/03/00
7.4	Execução dos treinamentos	28/02/00	13/11/00
7.4.1	Técnicos da CTD	28/02/00	17/03/00
7.4.2	Capacitação dos Coordenadores de Treinamento	23/03/00	24/03/00
7.4.3	G1 - Grupo Especial de Instrutores – Sede	20/03/00	05/04/00
7.4.4	G2 - Grupo Especial de Instrutores - Sede e Unidades Regionais	03/04/00	19/04/00
7.4.5	G3 - Servidores das Unidades Regionais	24/04/00	29/06/00
7.4.6	G4 - Agentes Censitários Municipais e Supervisores	15/05/00	13/07/00
7.4.7	Recenseadores	17/07/00	27/07/00
7.4.8	Avaliação	01/08/00	13/11/00
8	Diagramação final e impressão de instrumentos de coleta e administrativo	04/10/99	28/09/00
8.1	Diagramação final dos instrumentos de coleta	20/10/99	01/03/00
8.1.1	CD 1.01 - Questionário Básico	20/10/99	11/01/00
8.1.2	CD 1.02 - Questionário da Amostra	20/10/99	01/03/00
8.1.3	CD 1.03 - Folha de Domicílio Coletivo	20/10/99	03/12/99
8.1.4	CD 1.06 - Caderneta do Setor	20/10/99	03/12/99
8.1.5	CD 1.07 - Folha de Coleta	20/10/99	03/12/99
8.2	Impressão	04/10/99	20/07/00
8.2.1	CD 1.01 - Questionário Básico	15/05/00	30/06/00
8.2.2	CD 1.02 - Questionário da Amostra	04/10/99	20/07/00
8.2.3	CD 1.03 - Folha de Domicílio Coletivo	10/01/00	28/02/00
8.2.4	CD 1.06 - Caderneta do Setor	10/01/00	11/04/00
8.2.5	CD 1.07 - Folha de Coleta	26/06/00	10/07/00
8.3	Editoração e diagramação de documentos administrativos	11/02/00	16/06/00

Anexo 2.2 - Cronograma geral de atividades do Censo 2000
Versão Resumida

(continuação)

Id	Tarefas	Início	Término
8.3.1	AD 04 Ficha de Cadastro	14/02/00	15/03/00
8.3.2	AD 05 Ficha de Exclusão de Cadastro	14/02/00	15/03/00
8.3.3	AD 06 Requerimento de Dependência - Imposto de Renda	11/02/00	15/03/00
8.3.4	AD 07 Requerimento de Salário Família	14/02/00	15/03/00
8.3.5	AD 08 Declaração de Dependência Econômica	14/02/00	15/03/00
8.3.6	AD 09 Declaração Lei número 8.745	14/02/00	15/03/00
8.3.7	AD 10 Declaração de não acumulação de cargos públicos	14/02/00	15/03/00
8.3.8	AD 11 Programa de Assistência ao Censitário	14/02/00	15/03/00
8.3.9	AD 12 Contrato e termo de responsabilidade do Agente Censitário Municipal	14/02/00	15/03/00
8.3.10	AD 13 Contrato e termo de responsabilidade do supervisor	14/02/00	15/03/00
8.3.11	AD 14 Contrato e termo de responsabilidade do recenseador	14/02/00	15/03/00
8.3.12	AD 15 Crachá	14/02/00	15/03/00
8.3.13	AD 16 Ficha de contratação	14/02/00	15/03/00
8.3.14	AD 17 Ficha de Convocação	14/02/00	15/03/00
8.3.15	AD 18 Termo de desistência	14/02/00	15/03/00
8.3.16	AD 19 Requerimento de desligamento	14/02/00	15/03/00
8.3.17	AD 20 Comunicação de desligamento	14/02/00	15/03/00
8.3.18	AD 21 Ficha de Desligamento	11/02/00	15/03/00
8.3.19	AD 22 Pensão Alimentícia	11/02/00	15/03/00
8.3.20	AD 23 Solicitação de Auxílio Transporte Recenseador	11/02/00	15/03/00
8.3.21	AD 24 Solicitação do Auxílio - Transporte Mensalista	11/02/00	15/03/00
8.3.22	AD 25 Folha de Frequencia	11/02/00	15/03/00
8.3.23	AD 26 Folha de Afastamento	11/02/00	15/03/00
8.3.24	AD 27 Termo Aditivo do Agente Censitário Municipal e do Agente Censitário Supervisor	11/02/00	15/03/00
8.3.25	AD 28 Prorrogações	14/02/00	15/03/00
8.3.26	AD 29 Pagamento do adiantamento da produção	14/02/00	15/03/00
8.3.27	AD 30 Solicitação de pagamento por produção	14/02/00	15/03/00
8.3.28	AD 31 Solicitação de pagamento de ajuda de custo de transporte	14/02/00	15/03/00
8.3.29	AD 32 Declaração de desistência da funçaõ de Agente Censitário Municipal	14/02/00	15/03/00
8.3.30	AD 33 Contrato e termo de responsabilidade do ACA	14/02/00	15/03/00
8.3.31	AD 34 Contrato e termo de responsabilidade ANC	14/02/00	15/03/00
8.3.32	AD 36 Declaração de opção Edital AO2 2000	14/04/00	17/04/00
8.3.33	AD 37 Ficha de Cadastro Complementar	14/04/00	17/04/00
8.3.34	AD 38 Orientações Gerais Pessoal Mensalista SIAPE	14/04/00	17/04/00
8.3.35	AD 39 Contrato e Termo de responsabilidade do Agente de Pesquisa	14/04/00	17/04/00
8.3.36	AD 40 Orientações Gerais - Pessoal Mensalista ACM e ACS	14/04/00	12/05/00
8.3.37	AD 41 Declaração	14/04/00	12/05/00
8.3.38	AD 42 Declaração Retribuição do Recenseador	19/05/00	16/06/00
8.4	Impressão de documentos administrativos	20/03/00	10/07/00
8.4.1	AD 04 Ficha de Cadastro	20/03/00	10/04/00
8.4.2	AD 05 Ficha de Exclusão de Cadastro	20/03/00	10/04/00
8.4.3	AD 06 - Requerimento de Dependência - Imposto de Renda	20/03/00	10/04/00
8.4.4	AD 07 - Requerimento de Salário Família	20/03/00	10/04/00
8.4.5	AD 08 - Declaração de Dependência Econômica	20/03/00	10/04/00
8.4.6	AD 09 - Declaração Lei n. 8.745	20/03/00	10/04/00
8.4.7	AD 10 - Declaração de não acumulação de cargos públicos	20/03/00	10/04/00
8.4.8	AD 11 - Programa de Assistência ao Censitário	20/03/00	10/04/00
8.4.9	AD 12 - Contrato e termo de responsabilidade do Agente Censitário Municipal	20/03/00	16/05/00
8.4.10	AD 13 - Contrato e termo de responsabilidade do supervisor	20/03/00	10/04/00
8.4.11	AD 14 Contrato e Termo de Responsabilidade do Recenseador	02/05/00	19/05/00
8.4.12	AD 15 Crachá em branco - Analistas e Auxiliares	27/03/00	03/04/00
8.4.13	AD 15 - Crachá para o ACM e Supervisor	14/04/00	28/04/00
8.4.14	AD 15 - Crachá para o Recenseador	03/07/00	10/07/00

Anexo 2.2 - Cronograma geral de atividades do Censo 2000
Versão Resumida

(continuação)

Id	Tarefas	Início	Término
8.4.15	AD 16 - Ficha de Contratação	20/03/00	18/05/00
8.4.16	AD 17 - Ficha de Convocação	20/03/00	18/05/00
8.4.17	AD 18 - Termo de Desistência	20/03/00	18/05/00
8.4.18	AD 19 - Requerimento de Desligamento	20/03/00	18/05/00
8.4.19	AD 20 - Comunicação de Desligamento	20/03/00	18/05/00
8.4.20	AD 21 - Ficha de Desligamento	20/03/00	18/05/00
8.4.21	AD 22 - Pensão Alimentícia	20/03/00	10/04/00
8.4.22	AD 23 - Solicitação do Auxílio- Transporte Recenseador	20/03/00	18/05/00
8.4.23	AD 24 - Solicitação de Auxílio-Transporte Mensalista	20/03/00	10/04/00
8.4.24	AD 25 - Folha de Freqüência	20/03/00	18/05/00
8.4.25	AD 26 - Folha de Afastamento	20/03/00	10/04/00
8.4.26	AD 27 - Termo Aditivo (ACM / ACS)	20/03/00	19/05/00
8.4.27	AD 28 - Prorrogações	20/03/00	18/05/00
8.4.28	AD 29 -Pagamento do Aditamento da Produção	20/03/00	10/04/00
8.4.29	AD 30 - Solicitação de Pagamento por Produção	20/03/00	24/05/00
8.4.30	AD 31 Solicitação de Pagamento Ajuda de Custo Transporte	20/03/00	25/05/00
8.4.31	AD 32 - Declaração de desistência da função de Agente Censitário Municipal	20/03/00	10/04/00
8.4.32	AD 33 - Contrato e Termo de responsabilidade do ACA	20/03/00	10/04/00
8.4.33	AD 34 - Contrato e termo de responsabilidade ANC	20/03/00	10/04/00
8.4.34	AD 36 Declaração de opção Edital AO2 2000	14/04/00	17/04/00
8.4.35	AD 37 Ficha de Cadastro Complementar	14/04/00	17/04/00
8.4.36	AD 38 Orientações Gerais Pessoal Mensalista SIAPE	14/04/00	18/04/00
8.4.37	AD 39 Contrato e Termo de responsabilidade do Agente de Pesquisa	14/04/00	18/04/00
8.4.38	AD 40 Orientações Gerais - Pessoal Mensalista ACM e ACS	14/04/00	12/05/00
8.4.39	AD 41 Declaração	14/04/00	12/05/00
8.4.40	AD 42 Declaração Retribuição do Recenseador	19/05/00	16/06/00
8.5	Embalagem e expedição de material de coleta e administrativo	01/12/99	28/09/00
8.5.1	Embalagem	01/12/99	22/09/00
8.5.2	Expedição	03/03/00	28/09/00
9	Compras e distribuição de material	04/01/99	31/07/01
9.1	Compras e licitações nas Unidades Regionais	26/07/99	17/12/99
9.1.1	Definição de procedimentos administrativos	26/07/99	23/09/99
9.1.2	Elaboração de Manual de Recursos Materiais	27/09/99	26/10/99
9.1.3	Treinamento gerencial de compras e licitações p/ DERE e DIPEQ	09/08/99	17/12/99
9.2	Aquisição de mobiliário para as Unidades Regionais	04/01/99	29/06/00
9.2.1	Estudo e definição	04/01/99	26/08/99
9.2.2	Especificação técnica e quantificação	01/09/99	29/10/99
9.2.3	Editais, licitações e contratos	01/11/99	30/12/99
9.2.4	Entrega	03/03/00	09/06/00
9.2.5	Instalação	17/04/00	29/06/00
9.3	Aquisição de material para impressoras digitais	28/04/00	04/12/00
9.3.1	Elaboração do Edital	28/04/00	19/06/00
9.3.2	Análise e aprovação da PGE	19/06/00	10/07/00
9.3.3	Análise das propostas	10/07/00	13/10/00
9.3.4	Assinatura do contrato	16/10/00	27/10/00
9.3.5	Entrega	27/10/00	04/12/00
9.4	Compras e contratações de serviços para a gráfica	04/01/99	31/07/01
9.4.1	Preparo da área de armazenamento de papel	19/03/99	14/04/00
9.4.2	Aquisição de equipamentos para a gráfica	04/01/99	31/07/01
9.5	Compras de material para coleta	15/10/99	28/07/00
9.5.1	Bolsa do recenseador	15/10/99	28/07/00
9.5.2	Lapiseira e grafite	15/10/99	14/04/00
9.5.3	Prancheta para recenseador	22/11/99	14/07/00
9.5.4	Borracha para recenseador	15/10/99	12/04/00

Anexo 2.2 - Cronograma geral de atividades do Censo 2000
Versão Resumida

(continuação)

Id	Tarefas	Início	Término
9.5.5	Canetas esferográficas	15/10/99	28/04/00
9.5.6	Material de escritório	15/10/99	10/07/00
9.5.7	Envelopes para guarda de documentos	15/10/99	10/07/00
9.5.8	Papel para impressora	15/10/99	03/07/00
9.6	Compras de veículos	22/04/99	11/08/00
9.6.1	Etapa 1	22/04/99	14/12/99
9.6.2	Etapa 2	14/02/00	11/08/00
9.7	Sistema de Administração de Pessoal – SAP	01/11/99	30/06/00
9.7.1	Ajustes	01/11/99	28/01/00
9.7.2	Implantação para o Censo 2000	01/11/99	19/05/00
9.7.3	Treinamento Administrativo - ACM e ACS	22/05/00	30/06/00
9.8	Sistema Financeiro	01/06/99	14/07/00
9.8.1	Definição de procedimentos	26/07/99	23/09/99
9.8.2	Elaboração de Manual de Procedimentos Orçamentários e Financeiros	03/09/99	26/10/99
9.8.3	Levantamento da situação das Unidades Gestoras	01/06/99	30/06/99
9.8.4	Preparação das Unidades Gestoras	30/09/99	14/07/00
9.8.5	Solicitação autorização de Suprimento de Fundos com valor superior	01/12/99	30/12/99
9.8.6	Execução do Programa de Treinamento	03/04/00	28/04/00
9.9	Treinamento administrativo	01/05/00	11/05/00
9.9.1	Módulo Recursos Humanos- Coordenadores Administrativos e Assessores	02/05/00	08/05/00
9.9.2	Módulo Recursos Financeiros - Coordenadores Adm. E Assessores	01/05/00	11/05/00
9.10	Acompanhamento administrativo	16/03/00	30/04/01
10	Suporte tecnológico	05/05/98	15/01/01
10.1	Definição do modelo da Base de Dados	01/02/00	29/12/00
10.1.1	Planejamento do ambiente de dados para produção do Censo	01/02/00	06/03/00
10.1.2	Planejamento da capacidade para o ambiente de produção do Censo	06/03/00	28/04/00
10.1.3	Avaliação das alternativas para acesso	03/04/00	30/06/00
10.1.4	Avaliação das alternativas para o modelo	02/05/00	30/06/00
10.1.5	Definição do modelo	03/07/00	14/07/00
10.1.6	Definição dos procedimentos	17/07/00	18/08/00
10.1.7	Testes e ajustes	21/08/00	01/11/00
10.1.8	Documentação para produção	06/11/00	29/12/00
10.2	Projeto tecnológico	05/05/98	15/01/01
10.2.1	Upgrade do mainframe	02/10/98	01/08/00
10.2.2	Licitação da solução para entrada de dados	17/06/98	07/12/99
10.2.3	Aplicativo de entrada de dados do Censo Experimental	17/01/00	14/07/00
10.2.4	Execução da entrada de dados do Censo Experimental	17/07/00	28/07/00
10.2.5	Aplicativo de entrada de dados dos questionários do Censo 2000	31/07/00	15/01/01
10.2.6	Informatização das agências	31/03/00	05/01/01
10.2.7	Sistema de Indicadores Gerenciais da Coleta – SIGC	05/05/98	19/05/00
11	Execução da coleta de dados e pesquisa de avaliação	01/04/99	28/09/01
11.1	Postos de coleta	01/04/99	30/06/00
11.1.1	Estudos iniciais	01/04/99	25/08/99
11.1.2	Elaboração da proposta	01/04/99	28/05/99
11.1.3	Escolha dos locais	01/11/99	28/04/00
11.1.4	Instalações	28/03/00	30/06/00
11.2	Coleta de dados	14/02/00	30/10/00
11.2.1	Elaboração do plano de acompanhamento da coleta	14/02/00	30/06/00
11.2.2	Definição dos locais de acompanhamento da coleta	10/07/00	30/10/00
11.2.3	Formação das equipes para acompanhamento da coleta	01/08/00	30/10/00
11.2.4	Início da coleta	01/08/00	01/08/00
11.2.5	Coleta	01/08/00	15/12/00
11.2.6	Acompanhamento da coleta	08/08/00	15/12/03
11.3	Pesquisa de Avaliação – PA	31/08/99	28/09/01

Anexo 2.2 - Cronograma geral de atividades do Censo 2000
Versão Resumida

(continuação)

Id	Tarefas	Início	Término
11.3.2	Editoração dos instrumentos de coleta	10/04/00	31/07/00
11.3.3	Impressão dos instrumentos de coleta e material de treinamento	02/05/00	05/10/00
11.3.4	Comunicação às DIPEQs sobre os setores selecionados	09/10/00	09/10/00
11.3.5	Preparo dos mapas dos setores da PA (cópia dos mapas revisados)	16/10/00	03/11/00
11.3.6	Expedição do material	31/07/00	31/08/00
11.3.7	Treinamento	15/05/00	27/10/00
11.3.8	Execução da coleta de dados	16/10/00	28/02/01
11.3.9	Apuração	02/03/01	28/09/01
12	Apuração dos questionários do Censo 2000	17/06/98	08/01/03
12.1	Centros de Captura de Dados	17/06/98	01/11/01
12.1.1	Seleção dos coordenadores	18/08/98	03/04/00
12.1.2	Projeto tecnológico	17/06/98	15/11/00
12.1.3	Definição do Sistema de Empastamento do Censo 2000	01/03/00	18/10/00
12.1.4	Aquisição de caixa máster	04/10/00	02/11/00
12.1.5	Compras gerais de material	04/04/00	01/03/01
12.1.6	Aquisição dos scanners	21/07/00	16/11/00
12.1.7	Aquisição de no-break e estabilizadores	08/05/00	31/07/00
12.1.8	Aquisição de estabilizadores	03/05/00	31/07/00
12.1.9	Equipamentos de informática	21/07/00	30/10/00
12.1.10	Apuração	13/08/99	01/11/01
12.2	Desenvolvimento dos Sistemas de Produção	09/01/01	03/08/01
12.2.1	Consolidação dos dados enviados pelos CCDs - Questionários Básico e Amostra	09/01/01	09/03/01
12.2.2	Conjunto Universo: 1. Formação do Conj. / 2. Criação variáveis / 3. Carga BD	12/02/01	12/04/01
12.2.3	Formação dos lotes para imputação	20/02/01	30/04/01
12.2.4	Imputação do Conjunto Universo - Sistema DIA	20/02/01	07/05/01
12.2.5	Definição dos relatórios de acompanhamento e análise do universo	16/05/01	29/05/01
12.2.6	Imputação da Renda do Conjunto Universo	25/05/01	25/07/01
12.2.7	Plano de Análise da imputação do Conjunto do Universo	30/03/01	03/08/01
12.2.8	Carga banco dados imputados para o Conjunto Universo	13/02/01	18/06/01
12.2.9	Carga base de dados originais do Conjunto Amostra (sem variáveis derivadas)	09/01/01	09/03/01
12.2.10	Formação dos lotes do conjunto para codificação	12/02/01	12/04/01
12.2.11	Desenvolvimento do sistema de codificação automática / assistida	01/03/01	30/04/01
12.3	Sistema de Codificação	05/04/99	10/04/02
12.3.1	Definição do modelo (concepção do sistema)	05/04/99	27/09/99
12.3.2	Implantação na CTD	11/08/99	24/04/00
12.3.3	Execução da codificação	01/02/00	10/04/02
12.4	Sistema de Crítica	18/01/99	22/10/02
12.4.1	Planejamento da crítica de quantidade (Questionário Básico e Amostra)	01/03/00	05/03/01
12.4.2	Crítica do Conjunto Universo	03/01/00	14/11/01
12.4.3	Crítica do Conjunto Amostra	18/01/99	22/10/02
12.5	Carga no banco final	02/07/01	17/12/02
12.5.1	Definição das variáveis derivadas	02/07/01	13/07/01
12.5.2	Programação das variáveis derivadas	16/07/01	10/08/01
12.5.3	Geração do banco	09/10/02	17/12/02
12.6	Expansão da Amostra	01/03/99	18/10/02
12.6.1	Áreas de Ponderação e Disseminação	01/03/99	13/08/02
12.6.2	Estimação	03/01/00	18/10/02
12.7	Análise da qualidade socioeconômica e demográfica	19/09/02	08/01/03
12.7.1	Domicílios	19/09/02	08/01/03
12.7.2	Famílias	19/09/02	08/01/03
12.7.3	Nupcialidade	19/09/02	08/01/03
12.7.4	Fecundidade	19/09/02	08/01/03
12.7.5	Características gerais da população	19/09/02	08/01/03
12.7.6	Instrução	19/09/02	08/01/03

Anexo 2.2 - Cronograma geral de atividades do Censo 2000
Versão Resumida

(continuação)

Id	Tarefas	Início	Término
12.7.7	Migração	19/09/02	08/01/03
12.7.8	Mão-de-obra	19/09/02	08/01/03
13	Disseminação (divulgações)	07/10/99	20/01/04
13.1	Resultados Preliminares	30/10/00	21/12/00
13.1.1	Desenvolvimento dos textos	30/10/00	05/12/00
13.1.2	Desenvolvimento das tabelas	13/11/00	13/12/00
13.1.3	Desenvolvimento do CD-ROM	16/11/00	22/12/00
13.1.4	Inserção das informações na Internet	30/11/00	21/12/00
13.1.5	Inserção das informações no SIDRA	16/11/00	19/12/00
13.1.6	Desenvolvimento dos Cartogramas	16/11/00	14/12/00
13.1.7	Impressão e Acabamento do Volume	07/12/00	21/12/00
13.1.8	Distribuição para as Unidades Regionais	18/12/00	21/12/00
13.1.9	Divulgação dos Resultados Preliminares	21/12/00	21/12/00
13.2	Sinopse Preliminar	03/04/00	09/05/01
13.2.1	Desenvolvimento dos textos e gráficos	01/03/01	03/05/01
13.2.2	Desenvolvimento das tabelas (Brasil e UFs)	03/04/00	03/05/01
13.2.3	Desenvolvimento do CD-ROM (Brasil e UFs)	08/01/01	09/05/01
13.2.4	Inserção das informações na Internet	23/04/01	09/05/01
13.2.5	Inserção das informações no SIDRA	05/03/01	04/05/01
13.2.6	Desenvolvimento dos Cartogramas	26/03/01	24/04/01
13.2.7	Impressão e Acabamento do Volume	23/04/01	10/05/01
13.2.8	Distribuição para as Unidades Regionais	07/05/01	14/05/01
13.2.9	Divulgação da Sinopse	09/05/01	09/05/01
13.3	Resultados do Universo	07/10/99	21/01/02
13.3.1	Elaboração do Sumário	02/05/01	22/06/01
13.3.2	Plano Tabular	07/10/99	23/07/01
13.3.3	Tabulação Piloto	08/06/01	17/08/01
13.3.4	Tabulação Final dos Resultados do Universo	23/07/01	14/12/01
13.3.5	Desenvolvimento dos textos e gráficos	23/10/01	17/12/01
13.3.6	Banco de Dados Multidimensional	23/04/01	20/12/01
13.3.7	Desenvolvimento do CD-ROM (Brasil e UFs)	01/11/01	09/01/02
13.3.8	Inserção das informações na Internet	17/12/01	10/01/02
13.3.9	Inserção das informações no SIDRA	14/05/01	28/01/02
13.3.10	Desenvolvimento dos Cartogramas	04/06/01	06/12/01
13.3.11	Impressão e Acabamento do Volume	17/12/01	22/01/02
13.3.12	Distribuição para as Unidades Regionais	21/12/01	23/01/02
13.3.13	Divulgação dos Resultados do Universo	21/12/01	21/12/01
13.4	Tabulação Avançada	07/02/00	08/05/02
13.4.1	Planejamento	07/02/00	29/01/02
13.4.2	Preparo do lote	15/01/01	03/08/01
13.4.3	Execução da Codificação do Censo 2000 para as Tabulações Avançadas	25/06/01	23/11/01
13.4.4	Execução da Crítica do Censo 2000 para as Tabulações Avançadas	16/11/01	01/03/02
13.4.5	Programação	17/07/00	08/03/02
13.4.6	Tabulação	08/01/02	09/04/02
13.4.7	Elaboração dos textos, gráficos e cartogramas	08/02/02	22/04/02
13.4.8	Impressão e Acabamento do Volume	24/04/02	17/05/02
13.4.9	Desenvolvimento do CD-ROM	24/04/02	07/05/02
13.4.10	Inserção das informações na Internet	18/04/02	06/05/02
13.4.11	Inserção das informações no SIDRA	04/03/02	05/04/02
13.4.12	Distribuição para as Unidades Regionais	02/05/02	07/05/02
13.4.13	Divulgação da Tabulação Avançada	08/05/02	08/05/02
13.5	Resultados da Amostra	03/04/00	18/01/04
13.5.1	Primeiros Resultados da Amostra	24/04/02	20/12/02
13.5.2	Censo Demográfico 2000 Agregados p/ setor - resultados do universo	09/01/03	31/03/03

Anexo 2.2 - Cronograma geral de atividades do Censo 2000
Versão Resumida

(conclusão)

Id	Tarefas	Início	Término
13.5.3	Características Gerais da População	23/12/02	27/06/03
13.5.4	Migração e Deslocamento	28/01/03	03/07/03
13.5.5	Microdados da amostra "segunda edição"	01/10/02	16/06/03
13.5.6	Amostra de Uso Público	20/02/03	25/06/03
13.5.7	Base de Informação por área de ponderação	14/03/03	01/07/03
13.5.8	Educação	06/03/03	02/1203
13.5.9	Trabalho e Rendimento	21/05/03	07/01/04
13.5.10	Nupcialidade e Fecundidade	30/06/03	07/01/04
13.5.11	Famílias e Domicílios	02/05/01	02/01/04
13.5.12	Malha do Setor Censitário Rural Digital do Brasil – Situação 2000	10/09/02	25/06/03
13.5.13	Malha do Setor Censitário Rural Digital dos Municípios do Brasil	31/10/02	25/06/03
13.6	Metodologia do Censo Demográfico 2000	25/10/02	06/02/04
13.6.1	Elaboração da Proposta do Sumário	25/10/02	07/02/03
13.6.2	Aprovação pela CPO da Proposta do Sumário	17/02/03	17/02/03
13.6.3	Primeira Versão dos Capítulos	14/02/03	24/07/03
13.6.4	Leitura Crítica dos Capítulos	27/08/03	10/09/03
13.6.5	Segunda Versão dos capítulos	18/09/03	13/10/03
13.6.6	Redação Final	14/10/03	24/11/03
13.6.7	Editoração	25/11/03	08/12/03
13.6.8	Aprovação da editoração	11/12/03	16/12/03
13.6.9	Editoração final e envio para impressão	17/12/03	30/12/03
13.6.10	Impressão e Acabamento do Volume	31/12/03	20/01/04
13.6.11	Distribuição para as Unidades Regionais	14/01/04	20/01/04
13.6.12	Divulgação da Metodologia	15/01/04	15/01/04
13.7	Atlas do Censo Demográfico 2000	02/12/02	05/01/04

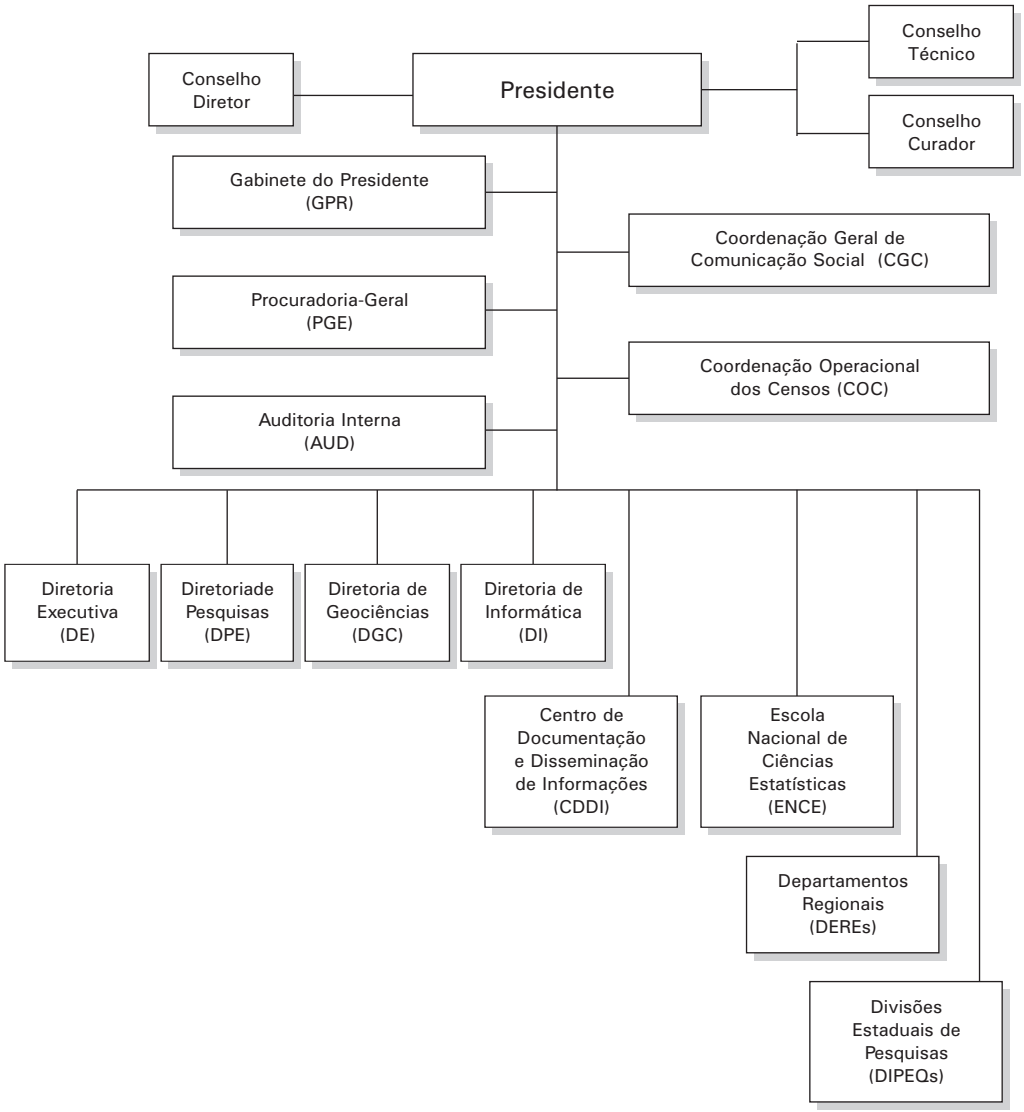
3 Logística da operação

A realização de um censo demográfico representa expressivo aumento da atividade interna do IBGE ao longo de cinco anos. São inúmeras tarefas que se desenvolvem simultânea e/ou seqüencialmente nas diversas áreas, concomitantes aos projetos e atividades contínuos da Instituição.

Assim, é preciso criar as condições básicas para que as atividades censitárias se desenvolvam de forma apropriada, sem causar prejuízos aos demais compromissos institucionais. Tais condições passam pelo redimensionamento do quadro de pessoal, da infra-estrutura e da organização gerencial, além de adaptações nas normas e nos sistemas administrativos, a fim de ajustá-los às especificidades de uma operação que se espalha por todo o território nacional e requer agilidade, controle e eficácia em todos os procedimentos operacionais, na gestão dos recursos humanos, orçamentários, financeiros e materiais.

Em vista do volume, da complexidade e do ritmo dos trabalhos censitários, é essencial a criação de estruturas paralelas à estrutura fixa do IBGE, com dedicação exclusiva ao Censo. Também se faz necessário o reforço das equipes permanentes que, adicionalmente a seus programas normais de trabalho, têm participação efetiva na atividade censitária. As instâncias, os órgãos e o quadro de pessoal temporários, criados para tratar do Censo 2000, serão abordados na primeira seção deste capítulo – Organização Geral. Contudo, é preciso notar que a estrutura fixa do IBGE, representada no organograma abaixo, é que suportou toda essa organização.

Figura 3.1 - Organograma geral do IBGE



Esta estrutura organizacional foi a que vigorou na maior parte do período da operação censitária, tendo sido implantada em janeiro de 2000, de acordo com o Estatuto do IBGE, aprovado pelo Decreto nº 3.272, de 3 de dezembro de 1999, e instituída pela Resolução do Conselho Diretor – CD nº 0015/99, de 7 de dezembro de 1999.

Este Capítulo também tratará de outros temas que, pelo aspecto inovador na Instituição e/ou pelas peculiaridades inerentes, demandaram atenção especial nas etapas de planejamento, preparo e implementação, quais sejam: as adaptações para a digitalização de documentos e reconhecimento ótico de caracteres, o desenvolvimento dos sistemas de controle e apuração, a montagem dos Centros de Captura de Dados – CCDs – bem como a impressão, distribuição e retorno de material.

3.1 Organização geral

Com o entendimento de que as condições ideais para a realização do Censo Demográfico 2000 somente seriam garantidas mediante entrosamento e ação conjunta de todas as áreas envolvidas, foi criada uma instância superior de planejamento, organização e acompanhamento geral da operação censitária – a Comissão de Planejamento e Organização Geral do Censo 2000 – CPO, que desempenhou papel fundamental desde o início do planejamento até o encerramento dos trabalhos do Censo. Para a implementação de cada uma das grandes etapas da operação censitária – ações preparatórias, coleta de dados, apuração e divulgação de resultados – o quadro de pessoal e a estrutura gerencial foram redimensionados, com a abertura de cerca de 235 mil vagas de pessoal temporário e a criação de aproximadamente 2 400 funções gratificadas. Às 530 agências permanentes do IBGE, responsáveis diretas pela execução da coleta dos dados, foram acrescentados mais 6 400 postos de coleta, abrangendo todos os municípios brasileiros.

3.1.1 Comissão de Planejamento e Organização Geral

A Comissão de Planejamento e Organização Geral do Censo 2000, instituída pela Resolução do Conselho Diretor nº 006/97, de 12 de junho de 1997, alterada pela Resolução do Conselho Diretor nº 020/99, de 30 de dezembro de 1999, foi presidida pelo diretor-executivo e composta pelos diretores, superintendentes e coordenadores dos comitês e das coordenações criados para gerir as atividades censitárias nas unidades centrais do IBGE. A CPO foi a instância superior de discussão de todas as questões relevantes, de orientação geral e acompanhamento da operação censitária. Coube-lhe, entre outras competências: estabelecer as diretrizes gerais para o planejamento, a organização e a execução da operação; decidir sobre questões estratégicas; determinar estudos e desenvolvimento de projetos; analisar e aprovar propostas e projetos apresentados pelas coordenações e grupos de trabalho; acompanhar a evolução das atividades; e apresentar ao Conselho Diretor propostas sobre assuntos de sua competência.

3.1.2 Coordenações e comitês das unidades centrais

No âmbito da Presidência do IBGE, atuou a Coordenação de Acompanhamento e Controle Operacional dos Censos – COC, unidade responsável pela logística da operação, incluindo o gerenciamento do orçamento e o acompanhamento do cronograma geral de atividades. A COC atuou também como Secretaria-Executiva da CPO e foi a responsável pelo gerenciamento dos processos seletivos para pessoal de coleta e pela elaboração e implementação, em conjunto com a Coordenação Técnica do Censo Demográfico 2000 – CTD do Programa de Treinamento do Censo 2000. Teve também sob sua subordinação administrativa os Centros de Captura de Dados, exceto o do Rio de Janeiro, que ficou subordinado à Diretoria de Informática.

Na Diretoria de Pesquisas – DPE, as ações foram conduzidas pelo Comitê do Censo 2000 da DPE e pela Coordenação Técnica do Censo Demográfico. Coube ao Comitê do Censo 2000 definir as diretrizes para o encaminhamento de todas as questões técnicas de competência da Diretoria de Pesquisas, como,

por exemplo: a definição do conteúdo dos questionários, dos conceitos e da metodologia de pesquisa, definição dos planos tabulares, análise e validação dos resultados; e a interação com o Censo Comum do Mercosul. Participaram, como membros desse comitê, o Diretor, o Diretor Adjunto, o Coordenador da CTD, os chefes dos Departamentos de População, de Emprego e Rendimento, de Metodologia e, ainda, técnicos especialistas nas áreas demográfica, socioeconômica e de informática.

A Coordenação Técnica do Censo Demográfico 2000 – CTD atuou como a unidade executiva e de coordenação de todas as atividades do Censo no âmbito da Diretoria de Pesquisas, tendo, entre suas atribuições: implementar as decisões tomadas no âmbito do Comitê do Censo 2000, da CPO e da Comissão Consultiva; gerenciar as etapas de planejamento, coleta e apuração das provas piloto, do Censo Experimental, do Censo 2000 e da Pesquisa de Avaliação da Cobertura da Coleta – PA, bem como preparar seus instrumentos de coleta e manuais técnicos e elaborar os planos de crítica e tratamento dos dados, de análise e de divulgação dos resultados.

Na Diretoria de Geociências – DGC, os trabalhos foram conduzidos pelo Comitê do Censo 2000 da DGC e pelas Coordenações Técnicas das Vertentes Urbana e Rural. A diretoria foi responsável por coordenar a preparação da Base Operacional e Geográfica, que envolveu: o inventário do mapeamento disponível – escala cadastral e topográfica – e de outros documentos necessários; o estabelecimento de parcerias para troca de materiais e consolidação das informações territoriais; a atualização e digitalização da documentação cartográfica; a elaboração de mapas de localidades, de mapas municipais e cadastros; a preparação da malha de setores censitários; o preparo da Caderneta do Setor; e a construção de uma base de referência para a divulgação dos resultados.

Na Diretoria de Informática – DI a Coordenação de Informática – COI orientou todas as ações relacionadas à operação do Censo 2000, realizadas nas diversas unidades da diretoria, e a Divisão de Sistemas de Censo – DICEN atuou no desenvolvimento e na implementação dos sistemas de controle e apuração do Censo. Entre as responsabilidades da área, destacam-se: a definição da tecnologia de suporte e a especificação das necessidades de equipamentos e treinamento, o desenvolvimento dos sistemas gerencial e de controle da coleta, de empastamento dos questionários, de captura de dados, de codificação, crítica e de imputação dos dados; e a montagem da rede nacional de comunicação do IBGE, interligando todas as agências da Instituição. Suas responsabilidades incluíram também a participação em definições dessa natureza para o Censo Comum do Mercosul.

Na Diretoria-Executiva – DE a preparação do Censo 2000 contou com a Coordenação de Apoio Administrativo ao Censo. Suas responsabilidades incluíram a coordenação das ações executadas nas diversas áreas da diretoria, entre elas: o recrutamento e seleção de parte do quadro temporário; a normalização das contratações de pessoal temporário e da gestão descentralizada do orçamento e da execução financeira, além do suprimento de recursos financeiros; o reaparelhamento das Unidades Regionais; a elaboração dos projetos e execução de obras; o preparo de áreas para armazenamento de material; as licitações e contratações dos serviços; a instalação dos CCDs; e a aquisição de mobiliário e equipamentos de informática, além dos materiais de consumo e outros, bem como da administração de sua distribuição.

No Centro de Documentação e Disseminação de Informações – CDDI a coordenação das atividades relativas ao Censo 2000 esteve a cargo da Coordenação de Atividades Gráficas e de Divulgação. Essas atividades incluíram: o reequipamento do parque gráfico; a aquisição das matérias-primas para impressão; a divulgação do Censo 2000, desde a definição da identidade visual até ações promocionais e de integração com a comunidade; a elaboração de documentos de divulgação; a impressão dos questionários, manuais, formulários e outros materiais; e o acabamento final e confecção dos produtos do Censo 2000, em meio impresso e digital. Couberam-lhe ainda a contratação das campanhas publicitárias de sustentação da coleta, a administração do *site* do Censo 2000, a colocação da base de dados do Mercosul na Internet, bem como a disseminação dos resultados.

A Escola Nacional de Ciências Estatísticas – ENCE, embora não tenha contado com uma coordenação específica para o Censo, participou, através do Departamento de Treinamento – DETRE do programa de treinamento, como responsável pelo preparo didático-pedagógico dos instrutores que participaram da cadeia de treinamentos técnicos.

Tal estrutura, liderada pela CPO e composta pelos comitês e coordenações, operou na sede do IBGE, no Rio de Janeiro, desde 1997, e foi complementada por coordenadores e supervisores que atuaram nas Unidades Regionais, estaduais e locais do IBGE. A seguir, estão listados os órgãos centrais de coordenação do Censo 2000.

Comissão de Planejamento e Organização Geral do Censo 2000 – CPO

Presidência
Coordenação de Acompanhamento e Controle Operacional dos Censos

Diretoria de Pesquisas
Comitê do Censo 2000
Coordenação Técnica do Censo Demográfico

Diretoria de Geociências
Comitê do Censo 2000
Coordenação da Vertente Urbana
Coordenação da Vertente Rural

Diretoria de Informática
Coordenação de Informática

Diretoria-Executiva
Coordenação do Apoio Administrativo ao Censo

Centro de Documentação e Disseminação de Informações
Coordenação de Atividades Gráficas e de Divulgação

3.1.3 Coordenação nas unidades regionais

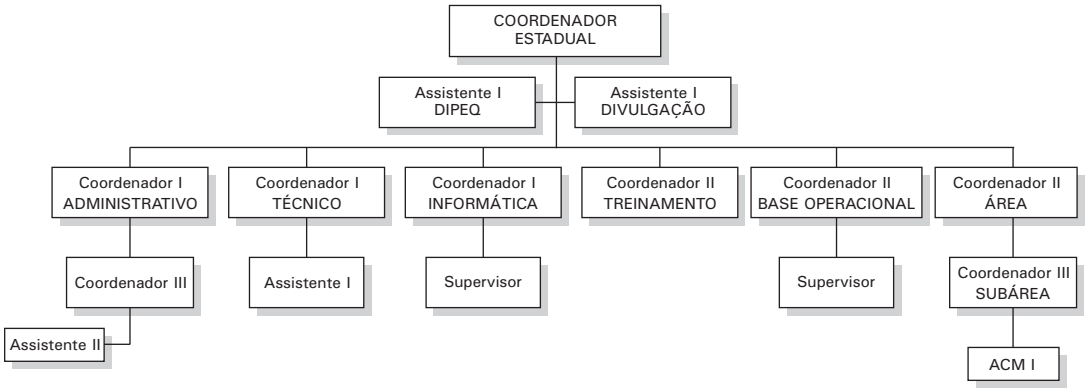
As Unidades Regionais foram responsáveis pela atualização da Base Operacional e Geográfica e pela coleta das informações, sob orientação técnica, respectivamente, das Diretorias de Geociências e de Pesquisas, recebendo orientação administrativa da Diretoria-Executiva. O controle operacional das atividades censitárias e a distribuição de recursos orçamentários foram realizados pela Coordenação de Acompanhamento e Controle Operacional dos Censos.

Nas Unidades Regionais, a estrutura organizacional permanente do IBGE contou com:

- 8 Departamentos Regionais – DEREs;
- 27 Divisões Estaduais de Pesquisa - DIPEQs;
- 5 Divisões de Geociências – DIGEOs; e
- 530 agências.

A essa estrutura foram acrescentados 6 400 postos de coleta, distribuídos em todos os municípios. O organograma a seguir apresenta, como exemplo, a estrutura de uma Divisão Estadual de Pesquisa de médio porte.

Figura 3.2 – Organograma de uma DIPEQ de médio porte



A unidade exemplificada – Divisão Estadual de Pesquisa de Santa Catarina – para executar o Censo, em seus 293 municípios e 6 833 setores, contou com:

- 21 agências;
- 315 postos de coleta;
- 315 Agentes Censitários Municipais – ACMs;
- 786 Agentes Censitários Supervisores – ACSs; e
- 6 329 Recenseadores.

Nas Unidades Regionais, a estrutura censitária obedeceu ao seguinte modelo:

- Coordenador Estadual – função acumulada pelo titular da Divisão Estadual de Pesquisa, como responsável por todas as atividades do Censo na Unidade da Federação;
- Coordenador Técnico – responsável pela coordenação de todo o trabalho técnico do Censo na Unidade da Federação, observadas as normas e procedimentos estabelecidos pela Diretoria de Pesquisas;
- Coordenador Administrativo – responsável pela coordenação de todas as atividades de cunho administrativo desenvolvidas na Unidade da Federação, atuando de acordo com as orientações da Diretoria-Executiva;
- Coordenador de Base Operacional e Geográfica – responsável pela supervisão técnica das tarefas de atualização da BOG na Unidade da Federação, recebendo orientações técnicas da Diretoria de Geociências;

- Coordenador de Informática – responsável pelas atividades de informática relativas às ações preparatórias e à execução do Censo no âmbito da Unidade da Federação, atendendo às normas e procedimentos estabelecidos pela Diretoria de Informática;
- Coordenador Estadual de Treinamento – responsável pela organização e supervisão operacional dos treinamentos censitários na Unidade da Federação, atuando como interlocutor da UR com a Coordenação Central do Treinamento do Censo 2000;
- Coordenador de Área – responsável pela orientação de todas as etapas da operação censitária em uma região – composta, em geral, pelas jurisdições de várias agências – estabelecendo a ligação entre os Coordenadores de Subárea e os Coordenadores Estaduais, Técnico e Administrativo, de forma a assegurar o adequado desenvolvimento das atividades censitárias na sua área de atuação;
- Coordenador de Subárea – responsável por todas as atividades técnicas, administrativas e de informática, nos municípios ou áreas subordinadas, que, em geral, correspondiam à jurisdição de uma agência;
- Assistente e Supervisor – atuaram junto aos coordenadores, oferecendo apoio e assessoramento em suas tarefas; e
- Agente Censitário Municipal – responsável pela organização do posto de coleta, pelo treinamento de Supervisores e Recenseadores, atuando de acordo com as instruções estabelecidas e sob a orientação do Coordenador de Subárea.

3.1.4 Estrutura gerencial

Para prover as unidades internas de uma estrutura gerencial capaz de suportar a carga de trabalho do Censo Demográfico 2000, foram criadas as Gratificações por Serviços Extraordinários – GSEs – concedidas na forma da Resolução do Conselho Diretor nº 003/2000, de 1º de março de 2000, com base no artigo 12 do Decreto-Lei nº 969, de 21 de dezembro de 1938, nas seguintes categorias e quantidades: Coordenador I (110), Coordenador II (359), Coordenador III (949), Assistente I (157), Assistente II (186), Supervisor (100) e Agente Censitário Municipal (572).

A estrutura gerencial censitária organizou e remunerou os servidores da Instituição que assumiram funções técnicas e administrativas, inclusive de gerenciamento do pessoal contratado. A supervisão técnica e o apoio administrativo da coleta envolveram cerca de 34 mil pessoas, das quais menos de 2,5 mil eram servidores do quadro permanente do IBGE. Os trabalhos de supervisão direta dos Recenseadores e de coordenação das atividades de cada posto de coleta foram funções para as quais o IBGE contratou e treinou pessoal temporário. Cada Agente Censitário Supervisor – ACS foi responsável por uma equipe de, em média, oito recenseadores, e cada posto de coleta ficou sob responsabilidade de um Agente Censitário Municipal – ACM. Essas equipes foram coordenadas por servidores da casa, com a seguinte estruturação: cada conjunto de, em média, oito municípios coube a um Coordenador de Subárea, e cada região com cerca de três subáreas foi supervisionada por um Coordenador de Área. Nas sedes das unidades regionais – Divisões Estaduais de Pesquisa e Departa-

mentos Regionais foram constituídas as coordenações técnica, administrativa, de treinamento e de informática, responsáveis pela coordenação e suporte à operação em nível estadual.

3.1.5 Recursos humanos

Uma operação censitária mobiliza grande contingente de pessoal, especialmente para a execução da etapa de coleta de dados. No Brasil, em função da legislação que rege o ingresso de pessoal nos órgãos públicos federais e do modelo de coleta adotado – que se prolonga por cerca de três meses – a mão-de-obra necessária tem de ser suprida com contratações temporárias, recrutada em processos seletivos simplificados.

As atividades censitárias preliminares, iniciadas dois anos antes da realização do Censo, determinaram o início das contratações já em 1998. No ano de 2000, para a coleta de dados, a incorporação de pessoal foi intensificada e estendeu-se até o ano de 2001, para atender à etapa de apuração dos dados coletados.

Durante o planejamento geral da operação, com base na experiência de censos anteriores e a partir da definição de alguns aspectos estratégicos, a estrutura de pessoal necessário pôde ser desenhada, com seus cargos, atribuições, prazos de contratação e perfis profissionais. Para atender a essa estrutura, sete processos seletivos foram realizados ao longo de toda a operação censitária.

3.1.5.1 Recrutamento e seleção

Os processos de recrutamento e seleção foram realizados de acordo com o Cronograma Geral de Atividades do Censo 2000 e em conformidade com os instrumentos legais, tais como a Lei nº 8.745, de 9 de dezembro de 1993; Lei nº 9.849, de 26 de outubro de 1999; Decreto nº 2.424, de 17 de dezembro de 1997 e Medida Provisória nº 1.748-40, de 8 de abril de 1999, e suas reedições, bem como nos termos do Decreto nº 3.298, de 20 de dezembro de 1999, que regulamenta a Lei nº 7.853, de 24 de outubro de 1989, que dispõe sobre a política nacional para a integração da pessoa portadora de deficiência, consolidando normas de proteção e dando outras providências.

A autorização para a realização do primeiro processo seletivo – para atender aos trabalhos de revisão da Base Operacional e Geográfica – deu-se pelo Despacho do Ministro de Estado da Administração Federal e Reforma do Estado, de 10 de março de 1998, publicado no Diário Oficial da União, de 11 de março de 1998. A autorização da seleção de pessoal para o Censo Experimental foi dada pelo Ofício nº 283 SEAP/MPOG, de 11 de maio de 1999, da Secretaria de Estado da Administração e do Patrimônio.

Os demais processos seletivos foram autorizados pela Portaria nº 127, de 21 de setembro de 1999, publicada no Diário Oficial da União, em 22 de setembro de 1999, do Ministro de Estado do Planejamento, Orçamento e Gestão. Para a obtenção de tal autorização, o IBGE encaminhou ofício ao Ministério justificando a necessidade das contratações, apresentando os quantitativos necessários para as diversas funções e seus respectivos prazos de contratação. Na oportunidade, informou também a existência dos recursos orçamentários exigidos. Os editais dos processos seletivos simplificados foram publicados no Diário Oficial da União, cumprindo o rito legal para a abertura de cada processo.

A seguir, são apresentados, por etapa da operação censitária a que atenderam, todos os Processos Seletivos Simplificados – PSSs realizados, cujos editais são apresentados no CD-Rom encartado nesta publicação.

a) Pessoal para a revisão da Base Operacional e Geográfica

1º PSS – Edital 01/98
Função: Auxiliar de Mapeamento Censitário – AMC

O processo seletivo para contratação de Auxiliar de Mapeamento Censitário visou à seleção de candidatos para atendimento às tarefas de atualização de mapas municipais, mapas cadastrais e croquis dos setores censitários da Base Operacional e Geográfica. Os candidatos a essa função, dos quais foi exigida a conclusão do ensino médio, submeteram-se a provas escritas objetivas de língua portuguesa, matemática e geografia e, ainda, à prova prática sobre conteúdo informativo constante do Manual do Candidato, com vistas à avaliação das habilidades para desenho, caligrafia e percepção de detalhes. A contratação dos candidatos classificados foi efetivada em junho de 1998, pelo período de 24 meses, com retribuição mensal de R$ 360,00.

A Divisão de Gerenciamento de Recursos Humanos – DE/CRH/DIGER do IBGE, foi a unidade responsável pelo processo seletivo.

b) Pessoal para o Censo Experimental

2º PSS – Edital 01/99
Funções: Agente Censitário Municipal – ACM, Agente Censitário Supervisor – ACS e Recenseador

O processo seletivo para o Censo Experimental foi realizado nos municípios de Marília/SP e de Bonito/PA, com vistas ao preenchimento de 29 vagas de Agente Censitário Municipal – ACM e Agente Censitário Supervisor – ACS bem como de 208 vagas de Recenseador, distribuídas entre os dois municípios. Para Marília/SP, foram reservadas 25 vagas de ACS e, para Bonito/PA, uma de ACM e duas de ACS. No caso da função de Recenseador, as 208 vagas foram preenchidas nos dois municípios, sendo 196 em Marília/SP e 12 em Bonito/PA.

A escolaridade exigida para a candidatura às vagas foi a conclusão do ensino médio para as funções de ACM e ACS e a conclusão do ensino fundamental para a função de Recenseador. Os candidatos às duas primeiras funções submeteram-se a provas escritas objetivas de língua portuguesa, matemática, geografia e conhecimentos específicos. Para a função de Recenseador, além das provas escritas objetivas, foi também aplicado um treinamento específico, cujo teste final tinha caráter eliminatório.

A prova de conhecimentos específicos foi baseada no conteúdo informativo constante da apostila do Censo Demográfico 2000. Para a função de ACM, além da aplicação das provas escritas objetivas, foi realizada entrevista técnica.

Os candidatos classificados nas três funções foram contratados pelo prazo máximo de cinco meses, a contar de agosto de 1999. A remuneração mensal foi diferenciada por função e por município, considerando os valores de mercado local e a tabela de salários do IBGE, cabendo aos ACMs a quantia de R$ 440,00, aos ACSs a quantia de R$ 360,00 e aos Recenseadores quantias variáveis de acordo com a produção apresentada – domicílios e pessoas recenseadas – no setor censitário trabalhado.

Aos candidatos à função de ACM, algumas atribuições e habilidades foram exigidas, a saber: organização de posto de coleta; domínio de conceitos e instruções técnico-administrativas; conhecimento da área de trabalho; coordenação,

acompanhamento e avaliação de desempenho dos ACSs; solução de problemas técnico-administrativos; organização e execução de treinamento dos Recenseadores; conferência dos instrumentos de controle da coleta e pagamento de Recenseadores; avaliação de novas necessidades de contratação de Recenseadores; transcrição de textos e confecção de planilhas em microcomputador; e participação na Comissão Censitária Municipal – CCM, quando determinada pelo superior.

Para a função de ACS, foram exigidas as seguintes especificações: domínio de conceitos e instruções técnico-administrativas; aplicação de treinamentos para Recenseadores; conhecimento de limites e características dos setores censitários sob sua responsabilidade; coordenação, organização, acompanhamento e avaliação do trabalho dos Recenseadores; realização de entrevistas, com preenchimento de formulários; preenchimento de formulários de controle da coleta; transcrição de dados em microcomputador; e realização das atividades administrativas do posto de coleta.

Aos candidatos à função de Recenseador exigiram-se os seguintes itens: domínio de conceitos e instruções de preenchimento dos instrumentos de coleta; conhecimento dos limites do setor censitário; realização de entrevistas; preenchimento de formulários dentro dos respectivos prazos; e visitas freqüentes ao posto de coleta, para apresentação do material coletado.

A Divisão de Gerenciamento de Recursos Humanos – DE/CRH/DIGER do IBGE, foi a unidade responsável pelo processo seletivo.

c) Pessoal para as atividades internas dos órgãos centrais

3º PSS – Edital 01/99
Função: Analista Censitário – AC

Seleção realizada para o preenchimento de 63 vagas, distribuídas por especialidades na cidade do Rio de Janeiro/RJ. Para o exercício de tal função, exigiram-se as seguintes atribuições: participação nas atividades relativas às etapas de planejamento, desenvolvimento e apuração dos dados do Censo; fomento das atividades de infra-estrutura; e desenvolvimento das atividades de disseminação dos dados. Neste processo seletivo, as vagas foram assim distribuídas: duas para Administração; trinta e duas para Análise de Sistemas; uma para Ciências Humanas; duas para Comunicação Visual; seis para Economia; dezesseis para Estatística; três para Programação Visual; e uma para Sociologia.

A seleção foi feita através de provas escritas objetivas de língua portuguesa, conhecimentos específicos e entrevista técnica, sendo esta última aplicada para todas as especialidades.

Os candidatos que fizeram a opção por concorrer à vaga da especialidade Ciências Humanas participaram de uma prova discursiva – prática em substituição à prova de conhecimentos específicos, e deveriam ter habilitação superior com formação voltada para essa área de conhecimento. Para as demais especialidades, foi requerido nível superior completo, correspondente à função a ser desempenhada.

Em relação ao salário, foi fixado o valor de R$ 1.100,00, a partir da contratação em novembro de 1999, durante o período de 24 meses.

O Núcleo de Computação Eletrônica – NCE da Universidade Federal do Rio de Janeiro – UFRJ, foi o responsável técnico pelo processo seletivo, coordenado pela DE/CRH/DIGER, sendo chamados os candidatos aprovados e classificados – em número cinco vezes maior ao número de vagas – para entrevista com técnicos do IBGE, escolhidos de acordo com a especialidade da unidade em que iriam trabalhar.

d) Pessoal para as atividades administrativas das Unidades Regionais

4º PSS – Edital 02/2000

Funções: Analista Censitário Contador – ACC e Auxiliar Censitário Administrativo – ACA

O processo seletivo para ACC e ACA preencheu, respectivamente, 33 e 707 vagas em cada função, distribuídas na quase totalidade das 26 Unidades da Federação e no Distrito Federal. As vagas para a função de ACC não foram disponibilizadas no Amapá, em Roraima e em Tocantins, em virtude de não se constituírem em unidades gestoras.

Aos candidatos, foi solicitada a comprovação de conclusão de curso superior em Ciências Contábeis, para a função de Analista Censitário Contador, e a comprovação de conclusão do ensino médio, para a função de Auxiliar Censitário Administrativo. Os candidatos às vagas de ACC submeteram-se a provas escritas objetivas de língua portuguesa e conhecimentos específicos, e os candidatos às vagas de ACA realizaram provas escritas objetivas de língua portuguesa, matemática e noções de administração.

Os valores fixados para a remuneração mensal foram de R$ 1.150,22 para a função de ACC e de R$ 360,00 para a função de ACA. A vigência dos contratos nas duas modalidades foi de 24 meses, observada a partir de março de 2000.

A Divisão de Gerenciamento de Recursos Humanos – DE/CRH/DIGER do IBGE, foi a unidade responsável pelo processo seletivo.

e) Pessoal para a coleta de dados

5º PSS – Edital 01/2000

Funções: Agente Censitário Municipal – ACM e Agente Censitário Supervisor – ACS

6º PSS – Edital 03/2000

Função: Recenseador

Foram realizados, por meio de provas escritas objetivas, dois processos seletivos para o preenchimento das vagas das três funções ligadas diretamente à coleta de dados – ACM, ACS e Recenseador. Para a execução dessa tarefa, o IBGE contratou, por meio de processo licitatório regulado pela Lei nº 8.666, de 21 de junho de 1993, empresa especializada, a Fundação Cesgranrio.

A primeira prova foi realizada no dia 9 de abril de 2000, selecionando Agentes Censitários Municipais e Agentes Censitários Supervisores, funções essas encarregadas da supervisão e controle das atividades dos Recenseadores, além do apoio operacional e administrativo aos postos de coleta. As 6 408 vagas para a função de ACM e as 25 283 para a função de ACS foram distribuídas entre todos os 5 507 municípios do País. Aos candidatos, foi exigida a conclusão do ensino médio e oferecidas as remunerações mensais de R$ 550,00 e R$ 440,00, para ACM e ACS, respectivamente.

A segunda prova foi realizada em 28 de maio de 2000, e selecionou 189 937 Recenseadores, dos quais foi exigida a conclusão do ensino fundamental e oferecida remuneração variável, de acordo com a produção cumprida em cada setor censitário.

A seleção dos Recenseadores contou com uma característica adicional: convocou para treinamento um contingente superior ao número de vagas por município, da ordem de 10%, o que possibilitou margem maior de escolha, com base no teste de aproveitamento do treinamento, bem como a reposição dos desistentes.

As provas para a seleção dos ACMs e ACSs foram aplicadas em 1 702 municípios pólos e, para Recenseadores, em 3 241, todos definidos sob a orientação das Unidades Regionais do IBGE, considerando-se a localização dos municípios e a distribuição dos candidatos inscritos.

A Empresa Brasileira de Correios e Telégrafos – ECT foi contratada pelo IBGE para prestação dos seguintes serviços:

- captação das inscrições;
- distribuição dos cartões de confirmação de inscrição;
- informação dos locais de provas;
- entrega e recolhimento dos malotes de provas; e
- divulgação de resultados e informação dos locais de treinamento.

Os processos de seleção de pessoal para coleta de dados incorporaram várias inovações, a saber:

- a inclusão de questões de conhecimentos técnicos, objeto de apostila denominada "O Trabalho do Recenseador", entregue no ato da inscrição. As questões incluídas davam informações gerais sobre o Censo, bem como sobre conceitos e procedimentos fundamentais ao exercício das funções mais diretamente ligadas à coleta de dados. Desta forma, os candidatos aprovados ingressavam no treinamento já com conhecimentos sobre a natureza do trabalho a ser realizado;
- a inclusão de questões sobre situações gerenciais na prova para as funções de ACM e ACS. Com esse esforço de maior adequação do exame às tarefas a serem cumpridas, elevou-se significativamente o nível gerencial dos contratados quanto ao desempenho de tarefas de supervisão e controle das atividades dos Recenseadores; e
- a divisão dos municípios de grande porte, como São Paulo/SP e Rio de Janeiro/RJ, por exemplo, em áreas de trabalho formadas por agregação de bairros vizinhos, a fim de oferecer aos candidatos a opção de trabalhar em localidades de sua preferência. A reserva de vagas para inscrição de candidatos interessados em trabalhar em áreas normalmente problemáticas para a coleta de dados, como favelas, povoados distantes das sedes municipais, localidades rurais de difícil acesso e/ou percurso, etc., propiciando, assim, a contratação de moradores – pessoas já adaptadas às peculiaridades locais. Esse mecanismo contribuiu, não apenas, para a redução das desistências em razão de dificuldades de acesso, percurso e outros problemas que costumam desmotivar o Recenseador, mas também para a agilização e melhor qualidade da cobertura da coleta.

Quanto aos critérios de aprovação e classificação, foram considerados não eliminados os candidatos que obtiveram notas diferentes de zero no conjunto das disciplinas integrantes das provas escritas objetivas. O somatório da pontuação indicou a ordem de classificação dos candidatos.

Considerando que uma única prova foi aplicada para a seleção de candidatos às funções de ACM e ACS, ao candidato mais bem classificado no município ou área de trabalho foi dada a prerrogativa de optar pela função de ACM, restando aos demais classificados a investidura nas funções de ACS disponibilizadas no edital.

Os candidatos à função de Recenseador, aprovados e classificados na prova escrita objetiva, foram convocados para treinamento, em quantitativo igual ao número de vagas a que concorreram, acrescido de 10%. Ao final do treinamento, foram submetidos a teste de aproveitamento, de caráter eliminatório. Contudo, a escolha do primeiro setor de trabalho seguiu a ordem de classificação na prova escrita objetiva.

No decorrer da coleta de dados, constatada a necessidade de reposição de Recenseador, depois de esgotada a reserva do contingente aprovado no treinamento, novos candidatos podiam ser convocados, obedecida a classificação da prova escrita objetiva e observado o número de vagas para o município ou área de trabalho previsto no edital.

Em obediência ao Decreto nº 3.298, de 20 de dezembro de 1999, foi reservado o equivalente a 5% do total das vagas oferecidas, em cada um dos processos seletivos, para os candidatos portadores de necessidades especiais.

Em razão das peculiaridades do trabalho de campo inerentes à coleta de dados, foram descritas detalhadamente nos editais dos processos seletivos as exigências mínimas, para o desempenho das funções, relativas à capacidade dos candidatos no que se refere à locomoção, acuidade visual, comunicação e escrita, entre outras.

Para análise dos laudos médicos apresentados pelos candidatos que indicaram, no requerimento de inscrição, a condição de pessoa portadora de necessidades especiais, à luz da legislação específica e dos critérios definidos nos regulamentos dos processos seletivos, o IBGE contou com a assistência de uma equipe multiprofissional, composta por profissionais da Associação Brasileira Beneficente de Reabilitação – ABBR, do Instituto Benjamin Constant, do Instituto Nacional de Educação de Surdos – INES, da Fundação Cesgranrio e do próprio IBGE.

No processo seletivo para provimento das vagas de ACM e ACS, a equipe multiprofissional analisou 1 068 laudos médicos, enquanto no de Recenseador 1 316.

A Coordenação de Acompanhamento e Controle Operacional dos Censos – COC foi a unidade do IBGE responsável pelos dois processos seletivos.

f) Pessoal para a apuração dos dados

7º PSS – Edital 05/2000

Funções: Auxiliar Censitário de Apuração – ACP e Auxiliar Censitário de Informática – ACI

Este processo seletivo foi realizado visando ao preenchimento de 1 147 vagas de Auxiliar Censitário de Apuração e 94 vagas de Auxiliar Censitário de Informática, nos cinco Centros de Captura de Dados. O nível de escolaridade exigido foi o de ensino médio concluído, e os candidatos foram submetidos à prova escrita objetiva, com 50 questões de múltipla escolha, de português, matemática, conhecimentos gerais – apenas para ACP – e Informática, somente para ACI.

O Núcleo de Computação Eletrônica – NCE da Universidade Federal do Rio de Janeiro – UFRJ, foi responsável pela execução desse processo seletivo, sob coordenação da DE/CRH/DIGER.

Os Supervisores de Turno, Supervisores de Módulo e Supervisores de CPD submeteram-se a treinamento, de caráter obrigatório e classificatório, buscando avaliar as características específicas de profissional com perfil para gerenciamento de pessoas. Participaram dessa etapa os classificados na prova escrita objetiva para a função de Auxiliar Censitário de Informática. Ao final do

treinamento, foi realizada uma prova escrita discursiva, e o primeiro colocado foi, então, indicado para Supervisor de Turno. Concorreram às vagas de Supervisor de CPD os candidatos classificados na prova escrita discursiva, de acordo com o número de vagas determinado para cada município sede de CCD. Para a categoria de Supervisor de Módulo, foram selecionados os candidatos que, na referida prova, obtiveram classificação inferior à dos Supervisores de CPD.

O Quadro 3.1 a seguir apresenta um resumo de todos os processos seletivos realizados para o Censo 2000.

Quadro 3.1 - Recrutamento e seleção de pessoal temporário para o Censo Demográfico 2000

Processos seletivos e funções	Vagas	Número de inscritos	Escolaridade	Alocação	Salário (R$)	Início da contratação
Edital 01/1998 - base operacional						
Responsável: IBGE/DE/CRH/DIGER						
Auxiliar de mapeamento censitário - AMC	700	23 000	Ensino médio	Capitais	360,00	jun. 1998
Edital 01/1999 - censo experimental						
Responsável: IBGE/DE/CRH/DIGER						
Agente censitário municipal - ACM	1	843	Ensino médio	Bonito (PA) e Marília (SP)	440,00	ago. 1999
Agente censitário supervisor - ACS	29				360,00	
Recenseador	220		Ensino fundamental		Variável por produção	
Edital 01/1999 - unidades da sede						
Responsável: IBGE/DE/CRH/DIGER e NCE/UFRJ						
Analista censitário - AC	63	1 015	Nível superior	Rio de Janeiro/ RJ	1 100,00	nov. 1999
Edital A 02/2000 - unidades regionais						
Responsável: IBGE/DE/CRH/DIGER						
Auxiliar censitário administrativo - ACA	707	28 000	Ensino médio	Capitais	360,00	mar. 2000
Analista censitário contador - ACC	33		Nível superior		1 150,22	
Edital 01/2000 - coleta de dados - mensalistas						
Responsável: IBGE/PR/COC e Fundação Cesgranrio						
Agente censitário municipal - ACM	6 408	459 958	Ensino médio	5 507 municípios	550,00	maio 2000
Agente censitário supervisor - ACS	25 283				440,00	
Edital 03/2000 - coleta de dados - produção						
Responsável: IBGE/PR/COC e Fundação Cesgranrio						
Recenseador	189 937	792 274	Ensino fundamental	5 507 municípios	Variável por produção	ago. 2000
Edital B 05/2000 - centros de captura de dados						
Responsável: IBGE/DE/CRH/DIGER e NCE/UFRJ						
Auxiliar censitário de apuração - ACP	1 147	26 812	Ensino médio	Campina Grande/PB Campinas/SP Curitiba/PR Goiânia/GO Rio de Janeiro/RJ	440,00	nov. 2000
Auxiliar censitário de informática - ACI	94				De 500,00 a 700,00	

3.1.5.2 Contratação

A contratação de pessoal temporário para o Censo 2000, exceto Recenseadores, ACS e ACM, foi executada pelas áreas de recursos humanos das Unidades Regionais, pela Coordenação de Recursos Humanos da Diretoria-Executiva, no caso das contratações feitas para as unidades da administração central, e pelos próprios Centros de Captura de Dados, no caso de seu pessoal. Essas contratações utilizaram os procedimentos habitualmente empregados pelo IBGE na contratação de pessoal temporário. Apesar de já ser de conhecimento das áreas de recursos humanos esse tipo de procedimento, novos manuais foram elaborados para facilitar o treinamento e o trabalho das equipes.

No caso do pessoal diretamente ligado à coleta – Recenseadores, ACSs e ACMs – a contratação, como também o pagamento, necessitou do desenvolvimento de um sistema informatizado específico, o Sistema de Administração de Pessoal – SAP. A contratação dos Recenseadores, em sua grande maioria, foi feita ao longo de uma semana, logo após o treinamento. Em razão da quantidade e da distribuição em todos os municípios, já que mais de 189 mil vagas foram abertas, essa atividade foi executada nos postos de coleta e nas agências, utilizando-se de rotinas e manuais específicos.

A seguir, serão abordados os procedimentos referentes à contratação do pessoal de coleta, que foram diferentes dos habituais, e os critérios gerais relativos a prazo e reposição de pessoal.

3.1.5.3 Contratação do pessoal de coleta

As contratações dos Agentes Censitários Municipais e dos Agentes Censitários Supervisores foram feitas pelos servidores do IBGE nos municípios sede das agências ou nos postos de coleta dos municípios de trabalho dos contratados. No caso dos Recenseadores, as contratações foram realizadas pelos ACMs e ACSs, nas agências e postos de coleta de todos os municípios.

Alguns procedimentos foram desenvolvidos visando a acelerar o processo de contratação e permitir o rápido início das atividades censitárias. Em especial, cabe citar:

- foram enviados a cada unidade de contratação de pessoal temporário arquivos com as listas dos candidatos aprovados, fornecidos pelas instituições responsáveis pelos processos seletivos, contendo o número de identificação do candidato que, depois da contratação, passou a ser sua matrícula funcional. Esses arquivos continham, ainda, todas as informações que os candidatos haviam fornecido no ato da inscrição, tais como: nome, número do documento de identidade, endereço, etc.;

- junto com as listas de divulgação de resultados dos processos seletivos, foram divulgadas as datas e os locais em que os candidatos aprovados e classificados, munidos da documentação exigida nos editais, deveriam comparecer para a formalização da contratação, no caso dos mensalistas, ou para os treinamentos, no caso dos Recenseadores; e

- os Recenseadores, durante o treinamento, preencheram o formulário Ficha de Cadastro, contendo informações pessoais, como os números dos documentos, endereço, dados de dependentes, etc. Para agilizar o processo de contratação, os formulários eram imediatamente encaminhados à DIPEQ, agência ou posto de coleta informatizado, para inclusão das infor-

mações no Sistema de Administração de Pessoal. Como o sistema já possuía as informações cadastrais fornecidas pelos arquivos dos processos seletivos, cabia ao operador apenas proceder as alterações necessárias e a inclusão de informações complementares. Uma das vias dos formulários era encaminhada ao posto de coleta onde ficaria lotado o Recenseador.

Todos os documentos relativos à contratação foram mantidos nas agências ou postos de coleta, em envelopes individualizados, até o final da operação censitária, quando então foram encaminhados às correspondentes DIPEQs.

a) Prazos e prorrogações

Todos os contratos tiveram prazo inicial de 30 dias, objetivando avaliar a adequação do contratado à função. Após o período inicial, os contratos poderiam ser prorrogados outras vezes por 30 dias, enquanto o serviço do contratado fosse necessário. O contrato e suas prorrogações não poderiam exceder o total de 24 meses. As renovações foram registradas nos contratos com dois dias de antecedência de seus términos. Para algumas funções, excetuadas as diretamente ligadas à coleta – ACS e Recenseador – foram permitidas prorrogações contratuais por períodos superiores a 30 dias, como ocorreu com o pessoal de nível superior, por exemplo.

b) Reposições

Em todas as funções em que se utilizou pessoal temporário, na hipótese de necessidade de reposição de vaga, o responsável pela contratação procedia à convocação de novos candidatos, obedecida a estrita ordem de classificação, de acordo com as listas fornecidas pelas instituições responsáveis pelos processos seletivos.

Todos os procedimentos mencionados acima foram descritos em manuais específicos, que ofereciam aos responsáveis pelas contratações o roteiro detalhado de todas as ações que deveriam adotar.

3.1.5.4 Efetivos de pessoal contratado para a coleta de dados

Para as funções de ACM e ACS foram contratadas, em geral, as quantidades previstas nos editais dos processos seletivos e esses quadros foram mantidos ao longo do período de coleta, visto que a mão-de-obra mensalista geralmente não apresenta alta rotatividade. Contudo, o mesmo não ocorreu com os Recenseadores. Para esta função, apesar de terem sido abertas 189 934 vagas no processo seletivo – que contou com a inscrição de mais de 790 mil candidatos – e de terem participado do treinamento 189 270 pessoas, foram registradas apenas 162 938 contratações, sendo 157 870 delas no mês de agosto, conforme mostra a Tabela 3.1. Isso significa dizer que cerca de 14% dos treinandos desistiram antes da contratação.

Tabela 3.1 - Número de recenseadores contratados e com setores concluídos

Recenseadores	Total	Meses			
		Agosto	Setembro	Outubro	Novembro
Contratados	162 938	157 870	148 329	88 603	9 471
Com setores concluídos	144 787	140 584	136 541	84 692	9 268

Fonte: IBGE, Diretoria Executiva, Sistema de Administração de Pessoal - SAP.

Esses números são explicados pelas características do trabalho do Recenseador que, a princípio, desperta o interesse de estudantes e outros trabalhadores de jornada parcial por ser tarefa ocasional e sem exigência de disponibilidade de tempo integral. Contudo, já no treinamento ou logo no início dos trabalhos, revelam-se as dificuldades inerentes à tarefa: áreas de difícil acesso e/ou percurso, obstáculos nas abordagens a condomínios e domicílios das classes mais altas, dificuldades de encontrar os moradores em casa ou a simples falta de adaptação à função de Recenseador. Entre outras relacionadas às especificidades das diferentes regiões do país, estas são historicamente as principais causas das desistências dos Recenseadores e foram também as observadas no Censo 2000.

As variações nas quantidades de Recenseadores ativos ao longo do período de duração da coleta – agosto a novembro de 2000 – estão demonstradas na Tabela 3.2. Pode-se notar que o mês de agosto – início da coleta – contou com o maior número de Recenseadores contratados – 157 870. Observa-se também as diferenças entre os 162 938 contratados no período e os que concluíram, ao menos, um setor censitário – 144 787; ou seja, 11% dos Recenseadores desistiram ou foram afastados por interesse do IBGE, antes de encerrar a coleta em ao menos um setor.

Tabela 3.2 - Recenseadores ativos, por período, segundo as Unidades da Federação

Unidades da Federação	Recenseadores com contratato ativo					
	Meses					Contratados no período
	Agosto	Setembro	Outubro	Novembro	Dezembro	
Brasil	**157 870**	**148 329**	**88 603**	**9 471**	**107**	**162 938**
Rondônia	1 207	1 046	626	113	2	1 253
Acre	376	364	216	20	-	400
Amazonas	1 711	1 998	1 775	827	-	2 304
Roraima	252	242	205	134	-	291
Pará	4 652	4 551	3 224	1 089	3	4 866
Amapá	351	325	181	17	-	354
Tocantins	1 066	1 004	681	76	-	1 100
Maranhão	5 086	4 755	3 382	98	-	5 190
Piauí	2 988	2 712	1 749	247	1	2 997
Ceará	5 889	5 334	3 370	363	-	5 921
Rio Grande do Norte	2 289	2 178	1 604	94	-	2 311
Paraíba	3 483	3 371	2 167	2	-	3 524
Pernambuco	7 075	6 964	4 753	561	-	7 380
Alagoas	2 364	2 236	1 500	20	-	2 384
Sergipe	1 983	1 821	913	1	-	1 996
Bahia	12 362	12 279	8 978	1 638	4	12 827
Minas Gerais	17 507	16 872	10 325	445	-	17 986
Espírito Santo	2 716	2 493	1 281	-	-	2 834
Rio de Janeiro	16 022	13 886	3 017	-	-	16 305
São Paulo	32 276	30 307	18 239	2 203	-	33 493
Paraná	9 311	9 215	6 414	693	96	9 835
Santa Catarina	5 177	4 676	2 741	81	1	5 290
Rio Grande do Sul	11 556	10 095	5 056	270	-	11 684
Mato Grosso do Sul	1 742	1 607	1 011	111	-	1 774
Mato Grosso	1 901	1 765	1 159	134	-	1 942
Goiás	4 564	4 376	2 978	99	-	4 664
Distrito Federal	1 964	1 857	1 058	135	-	2 033

Fontes: IBGE, Diretoria de Informática, Sistema de Indicadores Gerenciais da Coleta - SIGC e Diretoria Executiva, Sistema de Administração de Pessoal - SAP.

É importante registrar que as diferenças entre os números ao longo do período não significam somente saídas de Recenseadores por desistência ou por término de setor, vez que novos Recenseadores foram contratados durante o período e outros trabalharam em mais de um setor. Conforme mostra a Tabela 3.3, dos 215 790 setores censitários com unidades recenseadas, 212 797 foram coletados por Recenseadores, 2 402 o foram por ACMs e ACSs e 591 por servidores do quadro efetivo do IBGE.

Tabela 3.3 - Setores recenseados, por executor do trabalho, segundo as Unidades da Federação

Unidades da Federação	Total	Responsável pela coleta		
		Servidor	ACM ou ACS	Recenseador
Brasil	**215 790**	**591**	**2 402**	**212 797**
Rondônia	1 996	20	-	1 976
Acre	555	-	115	440
Amazonas	3 317	-	42	3 275
Roraima	507	38	6	463
Pará	6 148	18	57	6 073
Amapá	463	1	24	438
Tocantins	1 367	-	46	1 321
Maranhão	6 447	46	31	6 370
Piauí	3 717	3	18	3 696
Ceará	7 965	1	111	7 853
Rio Grande do Norte	2 655	9	2	2 644
Paraíba	4 171	-	9	4 162
Pernambuco	8 558	4	14	8 540
Alagoas	2 610	-	-	2 610
Sergipe	2 222	-	-	2 222
Bahia	15 342	11	137	15 194
Minas Gerais	22 523	41	176	22 306
Espírito Santo	3 228	15	11	3 202
Rio de Janeiro	20 756	72	68	20 616
São Paulo	49 713	9	1 214	48 490
Paraná	13 060	7	97	12 956
Santa Catarina	6 833	3	49	6 781
Rio Grande do Sul	16 907	10	78	16 819
Mato Grosso do Sul	2 714	17	41	2 656
Mato Grosso	3 377	69	2	3 306
Goiás	5 966	18	45	5 903
Distrito Federal	2 673	179	9	2 485

Fontes: IBGE, Diretoria de Informática, Sistema de Indicadores Gerenciais da Coleta - SIGC e Diretoria Executiva, Sistema de Administração de Pessoal - SAP.

3.1.6 Sistema de Administração de Pessoal

Com a experiência de operações censitárias anteriores, em especial as de 1991 e 1996, e após a busca de alternativas externas que pudessem atender às necessidades específicas do Censo, optou-se pelo desenvolvimento de um sistema informatizado próprio para efetuar o pagamento dos Agentes Censitários Municipais, Agentes Censitários Supervisores e Recenseadores. Tal sistema deveria ter alguns requisitos considerados fundamentais para o desenrolar da atividade censitária: garantir que as informações de produção – número de pessoas, questionários, unidades visitadas, etc. – utilizadas para o pagamento dos Recenseadores, fossem as mesmas do Sistema de Indicadores Gerenciais

da Coleta – SIGC; permitir aos seus múltiplos usuários o acesso fácil, rápido e eficiente, de forma a garantir que as diversas instâncias gerenciais tivessem uma visão do que estava sendo pago; e reduzir ao menor possível, o tempo transcorrido entre o aceite de um setor pelo Agente Censitário Supervisor e o pagamento ao Recenseador, fonte de reclamações em operações anteriores.

Em 1998, foi constituído o Grupo de Trabalho para Assuntos de Recursos Humanos – GTRH composto por servidores que, de alguma forma, já haviam participado de outros censos, com a atribuição de desenvolver o Sistema de Administração de Pessoal do Censo Demográfico 2000 – SAP. Para subsidiar seus trabalhos, o GTRH fez várias pesquisas documentais e consultou o relatório de avaliação das Unidades Regionais sobre as operações de 1996.

3.1.6.1 Desenvolvimento do sistema

O desenvolvimento do SAP foi feito pela Tecnocoop Sistemas, empresa contratada pelo IBGE para esse fim, contando com o apoio permanente do GTRH e da Diretoria de Informática para sua implantação em todas as unidades do IBGE.

A primeira versão do sistema foi desenvolvida para ser utilizada no Censo Experimental, realizado em Bonito/PA e Marília/SP. Sua filosofia era baseada no gerenciamento de cadastros e pagamentos, o que definiu a necessidade de divisão do sistema em dois módulos: um de cadastro e outro de pagamento. Já nessa primeira versão, o SAP funcionou interligado ao SIGC, e utilizando-se da rede interna de comunicação em cada estado, entre os postos de coleta dos municípios e as áreas de recursos humanos e financeiros dos Departamentos Regionais e das Divisões Estaduais de Pesquisa, sediadas nas capitais.

Além do desenvolvimento do sistema em seus dois módulos principais – Módulo Cadastro – SAPCAD e Módulo Pagamento – SAPPAG foram também elaborados os documentos de contratação e preparados os manuais de instruções sobre contratações e pagamentos.

Durante todas as fases de desenvolvimento, implantação, execução e manutenção, houve interação com as várias áreas do IBGE. As Unidades Regionais, por meio de suas equipes de recursos humanos, agências e postos de coleta, foram responsáveis pela entrada dos dados cadastrais e pelo processamento dos pagamentos dos contratados mensalistas e dos Recenseadores. Saliente-se que os Departamentos Regionais ficaram com a responsabilidade dos pagamentos de todo o pessoal mensalista de estados sob sua jurisdição e dos Recenseadores do seu próprio estado – pagamento por produção de cada setor trabalhado. As Divisões Estaduais de Pesquisa, localizadas fora da Unidade da Federação de seu Departamento Regional, foram responsáveis pela entrada de dados cadastrais e pelos pagamentos de seus Recenseadores.

Um fator que contribuiu para o êxito da implantação do SAP foi a realização de treinamento específico sobre os aspectos administrativos e gerenciais da operação censitária, que foi executado junto com o treinamento técnico-operacional. No caso do SAP, foram quatro horas de duração, atingindo os Grupos 1, 2 e 4 da cadeia de treinamentos do Censo 2000 – mais detalhes podem ser vistos no Capítulo 8 – Treinamento. Os outros servidores das agências foram treinados no Grupo 3, bem como os Agentes Censitários Supervisores, treinados no Grupo 4, receberam a instrução sobre o SAP em serviço, nas suas unidades de lotação.

3.1.6.2 Integração do SAP com o SIGC

Para agilizar o processo de pagamento e desburocratizar os procedimentos necessários à sua execução, foi estabelecido que o SAP utilizaria as informações geradas pelo SIGC. Assim, quando um setor adquiria o *status* de "liberado" no SIGC, essa informação era captada pelo SAP e, imediatamente, tinham início os procedimentos de pagamento.

Esse mecanismo permitiu a eliminação total de formulários de entrada de dados para pagamento, com informações sobre a identificação do setor e do Recenseador, quantidades de pessoas e domicílios recenseados, de unidades visitadas, de questionários preenchidos, etc., dispensando todas as etapas de emissão, tramitação e digitação de formulários. Assim, evitaram-se os erros inerentes a esses processos e ainda houve redução dos custos operacionais.

A integração SAP-SIGC realizava-se no órgão central de informática, no Rio de Janeiro/RJ, onde se encontravam os bancos de dados do SIGC, os quais eram diariamente carregados e/ou atualizados com as informações prestadas por todos os postos informatizados. Após a atualização, os novos bancos de dados eram transmitidos para o SAP e, então, distribuídos para os respectivos estados, a fim de que estes pudessem processar os pagamentos dos setores liberados. Assim, os Departamentos Regionais e as Divisões Estaduais de Pesquisa tiveram a opção de efetuar diariamente o pagamento dos setores concluídos, com a emissão de uma ou mais folhas de pagamento. Isto fez com que fosse plenamente atingida a meta de agilização dos pagamentos dos Recenseadores. Em média, o tempo decorrido entre a liberação e o efetivo pagamento do setor foi de cinco dias.

3.1.6.3 Módulo de Cadastro

Os programas e bancos de dados para operação do Módulo Cadastro foram instalados nos equipamentos do órgão central de informática, no Rio de Janeiro/RJ, e eram acessados pelos seus múltiplos usuários por meio da Intranet. Sua finalidade principal era reunir e disponibilizar os dados cadastrais dos contratados, para atender aos procedimentos de contratação/dispensa e de pagamento. A carga de dados foi realizada de três formas distintas: a primeira, centralizada a partir dos dados provenientes dos processos seletivos, quando foram carregados os dados de matrícula, nome, número da identidade, número do CPF, etc.; na segunda, também centralizada, foram incluídas as tabelas de referência, como as de taxas de pagamento, as dos descontos do INSS e do Imposto de Renda, etc.; e na terceira carga, descentralizada nas agências, postos de coleta informatizados, DEREs e DIPEQs, por ocasião da contratação efetiva dos candidatos aprovados, quando eram introduzidos os dados complementares de cada contratado, de acordo com a documentação exigida nos editais dos processos seletivos.

Deste módulo, constaram também outras tabelas utilizadas pelas unidades pagadoras, tais como: Função, Lotação, Arquivo de Setores, Dados Bancários, Tarifas de Auxílio-Alimentação e Auxílio-Transporte, Calendário, etc.

No módulo SAPCAD, eram incluídos ainda dados que interferiam diretamente no pagamento: falta/impontualidade – apenas para os mensalistas – ajuda de custo para transporte e adiantamento da produção – para Recenseadores – desligamento e afastamento. Este módulo oferecia também opção de consulta sobre a tramitação do pagamento de cada setor.

Os bancos de dados centralizados do módulo SAPCAD eram replicados para o módulo SAPPAG instalado nos estados. Assim, qualquer alteração ou inclusão de dados feita no cadastro central era replicada nos bancos instalados nas DIPEQs e nos DEREs, diariamente ou por comando ao longo do dia, com a finalidade de manter a integridade dos dados no sistema de pagamento.

3.1.6.4 Módulo de Pagamento

Este módulo ficou instalado em servidor UNIX de cada Departamento Regional e Divisão Estadual de Pesquisa, responsáveis por pagamentos a mensalistas e/ou Recenseadores, conforme anteriormente já descrito.

O módulo SAPPAG possibilitava a geração de folhas normais de pagamento de Recenseadores e de mensalistas, de folhas de correção, de gratificação natalina de mensalistas e de pagamento de verbas rescisórias. Além disso, permitia a emissão de relatórios gerenciais de cada uma das folhas de pagamento emitidas.

A atualização das informações cadastrais necessárias aos pagamentos ocorria pela replicação dos dados do módulo SAPCAD – centralizado – para os servidores UNIX de cada DERE ou DIPEQ, mas eram permitidas também inclusões manuais de rubricas para pagamentos e informações, feitas diretamente pelos operadores locais.

O módulo SAPPAG recebia do módulo SAPCAD os dados cadastrais, as tabelas e outras informações necessárias ao pagamento e obtinha do SIGC as informações sobre as quantidades produzidas para cada item de pagamento. Além disso, o sistema dispunha de arquivos específicos, como o do Plano de Assistência do Censitário – PAC oferecido pela Sociedade Ibgeana de Assistência e Seguridade – SIAS. Para finalizar o pagamento, o módulo SAPPAG oferecia a opção de geração de arquivos magnéticos das folhas de pagamento, que eram então enviados ao Banco do Brasil, para efetivação ou repasse a outros bancos, quando se fazia necessário.

O Sistema dispunha ainda de opções de atualização e consulta às tabelas de rubricas, taxas e encargos, bem como de emissão de documentos para classificação contábil e de rotinas anuais, como Ficha Financeira, DIRF e Declaração de Rendimentos.

Como já destacado anteriormente, um dos fatores que mais contribuiu para a agilização e o baixo número de erros nos pagamentos dos Recenseadores, em relação a censos anteriores, foi a integração do SAP com o SIGC. De uma forma simplificada, essa integração pode ser descrita com o seguinte fluxo: sendo dada a coleta do setor como concluída e aceita pelo responsável – ACS, ACM ou Coordenador de Subárea – seus dados finais eram incluídos no SIGC. A seguir, eram automaticamente acionados os mecanismos de controle interno do SIGC para cálculo de um conjunto de indicadores. Se seus valores estivessem em conformidade com parâmetros estabelecidos estatisticamente, o setor assumia o *status* de liberado. Assim, o setor era disponibilizado para o SAP, com seus dados básicos necessários ao cálculo do pagamento, a saber: código do setor, matrícula do Recenseador, quantidades coletadas por item de pagamento e datas de início e término da coleta. No momento da importação desses dados, o SAP procedia à validação do código do setor e da matrícula do Recenseador, verificando se os mesmos constavam da sua base de dados. Uma vez esses dados verificados e aprovados, o cálculo era processado, utilizando-se as quantidades informadas e seus respectivos valores, segundo a taxa de pagamento predefinida para o setor. Uma opção de controle permitia a verificação dos setores importados do SIGC, possibilitando o acesso a todos os dados dos setores pagos.

A maior parte dos pagamentos foi realizada na rede bancária. No caso dos mensalistas, todos os pagamentos foram feitos em contas correntes individualizadas. Para os Recenseadores, foram oferecidas três opções, a saber: depósito em conta corrente individualizada; crédito disponibilizado em nome do Recenseador no Banco do Brasil, passível de saque em qualquer agência; e por suprimento de fundos, pagamento em dinheiro, feito por servidor do IBGE diretamente ao Recenseador, nos municípios e localidades que não dispunham de rede bancária. Neste caso, o módulo SAPPAG gerava o recibo a ser firmado pelo Recenseador.

3.1.7 Administração financeira

À época da realização do Censo 2000, para administração dos trabalhos de rotina do IBGE, apenas os Departamentos Regionais e as Divisões Estaduais de Pesquisa do Rio de Janeiro e do Rio Grande do Sul eram unidades gestoras de recursos orçamentários e financeiros, sendo as demais Divisões Estaduais de Pesquisa dependentes dos respectivos Departamentos Regionais na realização de processos de compra de bens e serviços, de locações e de pagamentos em geral, inclusive de pessoal – salários, diárias, passagens, etc..

Levando-se em conta a grande quantidade de pagamentos e o volume de recursos que seriam movimentados durante a operação censitária nas Unidades Regionais, bem como ser fundamental propiciar agilidade aos processos de pagamentos, a direção do IBGE, por recomendação da Comissão de Planejamento e Organização Geral do Censo 2000 – CPO, após ouvir os titulares das Divisões Estaduais de Pesquisa e dos Departamentos Regionais, resolveu estender a algumas outras DIPEQs a autonomia de gestão. Essa atividade foi desenvolvida pela Coordenação de Orçamento e Finanças – COF da Diretoria-Executiva, que ficou responsável pela capacitação e acompanhamento das novas unidades gestoras.

Todos os pagamentos do Censo 2000 foram feitos através do Sistema Integrado de Administração Financeira – SIAFI, como acontece em toda a administração pública federal.

Como, nas operações censitárias, é elevada a quantidade de despesas de pronto pagamento, que precisam ser quitadas em dinheiro, das quais são exemplos os pagamentos a barqueiros, a guias, bem como a compra de passagens e outros gastos de pequeno porte, exigidos nos trabalhos de campo em localidades sem rede bancária, foi desenvolvido e implantado um novo sistema para emissão e controle dos valores disponibilizados a servidores do IBGE, através da modalidade de suprimento de fundos. Esse mecanismo permite ao servidor receber em sua conta bancária e sacar em espécie um valor cuja aplicação é previamente definida, a fim de efetivar pagamentos diretos a contratados temporários, prestadores de serviço ou fornecedores.

Esta seção tratará do projeto de descentralização das unidades gestoras regionais e do novo Sistema de Suprimento de Fundos, que representaram, senão as maiores, as mais significativas alterações nas rotinas da área financeira no Censo 2000.

3.1.7.1 Descentralização das unidades gestoras regionais

Unidades gestoras são aqui entendidas como aquelas providas de recursos orçamentários e financeiros, com competência para executarem despesas. Regularmente, no IBGE, havia apenas dez Unidades Regionais com essa autonomia. Para o Censo 2000, pelas razões já expostas, mais 14 unidades recebe-

ram tal autonomia, propiciando agilidade aos procedimentos administrativos durante a execução da etapa de coleta de dados.

A descentralização abrangeu todas as Divisões Estaduais de Pesquisa, exceto as unidades do Amapá, Tocantins e Roraima que, pelo porte e pela falta de pessoal, optaram por continuar vinculadas aos Departamentos Regionais. Com isso, foi possível imprimir maior velocidade na disponibilização de créditos orçamentários e recursos financeiros, tornando mais curtos os prazos de tramitação dos processos de aquisição de bens e serviços e de pagamentos de pessoal. Tal medida favoreceu a execução dos trabalhos, vez que propiciou sensível redução no tempo decorrido entre a tomada de decisão e a efetivação do ato administrativo.

Para os Coordenadores Administrativos das 24 unidades gestoras regionais foi ministrado treinamento sobre "Recursos Orçamentários e Financeiros", sendo este um dos módulos do treinamento administrativo, parte integrante do Treinamento técnico-operacional do Censo 2000. Os contadores contratados temporariamente para as unidades que não dispunham desse profissional em seus quadros receberam também o treinamento, acrescido de palestras, entre outros temas, sobre disponibilidade orçamentária, empenho e liquidação de despesa, apropriação contábil, programação e execução financeiras, conciliação bancária, suprimento de fundos e atribuições dos contadores.

3.1.7.2 Suprimento de fundos

Foram utilizadas, no Censo 2000, as duas modalidades de pagamento rotineiramente adotadas pelo IBGE, quais sejam: por depósito em conta bancária do favorecido, para pagamentos efetuados a fornecedores de bens e serviços – pessoa física ou jurídica – e a servidores efetivos e temporários; e por suprimento de fundos, cujo valor é disponibilizado em favor de servidor do quadro permanente da Instituição, a fim de que este providencie as aquisições necessárias e efetue diretamente os pagamentos, em espécie ou em cheques de conta vinculada "tipo B" – aberta para este fim específico –, no prazo determinado pelo ordenador de despesa.

O IBGE, com o intuito de atender às necessidades da operação censitária, obteve autorização do Ministério do Planejamento Orçamento e Gestão para aumentar o valor do limite para concessão de suprimento de fundos, que passou de R$4.000,00 para R$ 10.000,00, durante o exercício fiscal de 2000. A autorização foi objeto da Portaria nº 125, de 15 de junho de 2000, publicada no Diário Oficial da União, de 16 de junho de 2000, e retificada em 21 de junho de 2000, do Ministro de Estado do Planejamento, Orçamento e Gestão, que autorizou "conceder Suprimentos de Fundos Especiais no valor máximo de R$10.000,00 (dez mil reais) por suprimento, com prazo de aplicação de até 60 (sessenta) dias cada, para atender despesas de pronto pagamento referentes aos trabalhos do Censo Demográfico 2000".

Dois pressupostos orientaram o desenvolvimento de um sistema informatizado que tinha o propósito de alcançar plena efetividade nos processos de concessão e nos controles dos pagamentos feitos por suprimento de fundos. Foram eles: definir a ferramenta de trabalho a ser utilizada por todas as unidades gestoras do IBGE; e tornar disponíveis aos gestores, em tempo real, dados suficientes que propiciassem o adequado gerenciamento das informações.

Assim, a primeira abordagem do tema teve a finalidade de atualizar a norma interna que tratava do assunto, com vistas ao alcance dos resultados esperados e, bem como, à realidade tecnológica do IBGE, respeitados todos os princípios básicos impostos pela legislação vigente. Um outro aspecto não menos relevante foi o de proporcionar, a um só tempo, maior agilidade e confiabilidade ao processo de trabalho, já que as atividades desenvolvidas durante os censos determinam maior velocidade e exigem segurança dos gestores, em suas decisões.

O Sistema de Suprimento de Fundos foi concebido e desenvolvido a partir da preocupação em uniformizar os processos de trabalho, sendo realizado em parceria com algumas Unidades Regionais e com a Diretoria de Informática. A parceria com as unidades gestoras regionais serviu de base para a definição dos aspectos operacionais de registro, controle e amplitude do sistema, enquanto a participação da DI favoreceu o aspecto da sua concepção lógica. O sistema foi testado durante a realização do Censo Experimental e implantado definitivamente em cada uma das unidades gestoras regionais, a partir do início do Censo 2000, com administração local e base de dados residentes no servidor de rede de cada unidade.

A Diretoria de Informática instalou o sistema nas Unidades Regionais, em conjunto com os Coordenadores de Informática locais, ficando com a incumbência de dar suporte nos aspectos relativos a *hardware* e de transferir diariamente para a Coordenação de Orçamento e Finanças – COF cópias das bases de dados atualizadas.

Do ponto de vista do desenvolvimento das atividades de requisição, concessão, comprovação, aprovação e contabilização para suprimento de fundos, o sistema informatizado possibilitou uma visão completa da realização da despesa na unidade gestora e em suas unidades administrativas subordinadas, retratando a natureza da despesa realizada. Tais ações permitiram à COF acompanhar a movimentação realizada pelas unidades, avaliando procedimentos, criticando registros operacionais, indicando soluções e orientando as tomadas de decisão. Problemas mais complexos eram sanados no momento da visita do Supervisor Administrativo. As conclusões resultantes desse acompanhamento serviram para a revisão dos manuais operacionais e também para a adoção de providências imediatas quanto às novas orientações, a fim de prever e corrigir eventuais erros.

3.2 Escolha da tecnologia de captura de dados

O Censo Demográfico 1991 utilizou a digitação dos questionários como forma de captura dos dados. Para a entrada de dados do Questionário Básico, foi feito um estudo quanto à alternativa do uso de leitoras óticas, não sendo recomendado à época a sua utilização, pois os equipamentos existentes advinham de tecnologia da década de 80, já em obsolescência, que apresentava sérias restrições quanto ao seu uso em grande escala. Nesse Censo, a atividade de entrada de dados, assim como outras etapas da apuração, foi descentralizada em 21 centros regionais, sendo a definição do número de centros de apuração baseada no volume de questionários previstos.

A Contagem da População 1996 teve todo o seu encaminhamento voltado para o uso de leitoras óticas de marcas, visando dar maior agilidade à apuração. O modelo de questionário utilizado foi elaborado em tamanho A4, com as características específicas requeridas por essa tecnologia. Na coleta, esse tamanho de formulário foi utilizado pela primeira vez em censos populacionais, e não houve relato de inconvenientes quanto ao seu manuseio. As dificuldades encontradas referiam-se ao preenchimento do documento, principalmente no que toca ao hachuramento dos campos de marcas, que exigiu atenção e cuidados do Recenseador.

Manteve-se a filosofia da descentralização da atividade, com a utilização de 13 centros de apuração. O uso dessa tecnologia permitiu a redução de tempo e custos, se comparados com o processo de digitação convencional.

No final da década de 1990, presenciou-se um grande avanço na tecnologia de reconhecimento ótico, com o uso de imagens digitalizadas para o processamento de documentos. Investimentos foram feitos nessa área em termos de *software* e componentes periféricos – *scanner*, placas para melhoramento da imagem, mídias com grande capacidade de armazenamento – permitindo uma confiabilidade cada vez maior dessa tecnologia.

É importante salientar que a introdução de novas tecnologias deve ser feita após a realização de testes e somente quando as vantagens sejam claramente visíveis, e os riscos associados suficientemente previsíveis, controláveis e aceitáveis.

As experiências que vinham sendo realizadas por diversos países, se não permitiam afirmar, pelo menos indicavam que a tecnologia de reconhecimento ótico de marcas e caracteres era a melhor alternativa para o processamento de grandes volumes de documentos, pois se mostrava superior, mais flexível e confiável do que a tecnologia das leitoras óticas, além de infinitamente melhor do que a digitação tradicional.

Essa tecnologia permite reconhecer código de barras, marcas, números impressos, números e textos manuscritos, com maior flexibilidade quanto aos aspectos de impressão, corte e preenchimento dos formulários do que o requerido pelas leitoras óticas.

Dando início ao estudo dessa tecnologia, o IBGE realizou a Primeira Prova Piloto do Censo Demográfico 2000. Os resultados obtidos mostraram ser perfeitamente viável e confiável o seu uso no Censo, e as adaptações que foram necessárias nos formulários não causaram nenhum impacto ou prejuízo à obtenção das informações durante a coleta.

Houve a visita de três técnicos da Instituição aos Estados Unidos – ao U.S. Census Bureau e ao centro de digitalização da Receita Federal de Massachussets – para conhecerem, de perto, a utilização dessa tecnologia. Também foi efetuada a visita de um técnico ao Instituto Nacional de Estadística y Censos – INDEC, da Argentina, para o acompanhamento da Terceira Prova Piloto de seu Censo, em fase de processamento.

O passo seguinte foi o corpo técnico do IBGE adquirir conhecimento sobre os diversos aspectos relativos ao processamento de questionários, utilizando a nova tecnologia, tais como:

- obtenção de um modelo gráfico dos questionários que satisfizesse, simultaneamente, aos requisitos da coleta do Censo – compreensão, leitura, preenchimento, manuseio, etc. – e do reconhecimento ótico;
- avaliação das taxas de reconhecimento e das taxas de erros observadas para cada tipo de campo dos questionários – marcas, numéricos e alfanuméricos – e para os caracteres, no caso do preenchimento de textos;
- verificação das velocidades de processamento alcançadas em cada uma das etapas do processo de captura;
- avaliação das configurações de equipamentos – *hardware* – necessários ao processo;

- avaliação do fluxo operacional do trabalho – *workflow* - incluindo o controle de qualidade; e

- avaliação do nível de informatização dos processos.

O objetivo global dessa etapa foi dotar o IBGE da capacidade de definir quanto ao uso ou não da tecnologia, além conhecer com precisão o alcance e as características dos produtos disponíveis, visando criar conhecimento interno que possibilitasse a definição de requisitos necessários para a contratação de serviços, equipamentos e *softwares*, bem como decidir quantos Centros de Captura de Dados – CCDs – seriam necessários para atender ao prazo estipulado para o processamento.

A partir desse ponto, foram estabelecidas as estratégias para a entrada de dados do Censo 2000, as quais estão relatadas a seguir.

O desenvolvimento do sistema para o processamento das imagens e captura dos dados foi contratado à empresa especializada Politec Ltda, através de edital público, pois o conhecimento requerido para execução da tarefa não era de domínio dos técnicos da Instituição, e não pareciam convenientes o investimento e o tempo necessário para a obtenção interna dessa capacidade.

Uma questão, relacionada ao reconhecimento de textos manuscritos, exige um nível mínimo na qualidade da escrita. Portanto, a solução técnica a ser adotada devia garantir que a informação pudesse ser interpretada a partir dos registros nos questionários, no estado em que se achassem. Outro pressuposto requerido foi a possibilidade de recuperação das imagens digitalizadas de todos os documentos – questionários e Folhas de Coleta.

No processo de reconhecimento, com interpretação de marcas, números e caracteres manuscritos, existem sempre erros – de não reconhecimento e de falso reconhecimento – inerentes à sua implementação,

Como é necessário manter esses erros em níveis aceitáveis, torna-se importantíssimo o estabelecimento de um sistema de controle de qualidade, que possibilite o monitoramento desse processo de entrada de dados. Deve-se estabelecer um padrão de qualidade aceitável para o reconhecimento da informação de cada tipo de campo do questionário, buscando um equilíbrio entre qualidade desejada e eficiência operacional do processo. Por exemplo, os campos de marcas, por serem menos sujeitos a problemas de interpretação, devem ter taxas de erro aceitáveis mais rigorosas.

Outro ponto importante discutido foi quanto à operacionalização da atividade de entrada de dados. Algumas das alternativas estudadas, correspondiam a centros de captura complexos, com grande quantidade e diversidade de equipamentos, razoáveis contingentes de pessoas, levando a uma administração do trabalho onerosa e complicada.

O modelo adotado nos últimos censos, fez com que o IBGE definisse os equipamentos necessários, fizesse sua aquisição – incluindo instalação e manutenção – contratasse mão-de-obra temporária para os serviços de operação, arcando com a sua administração; isso aconteceu com o processo de leitura ótica usado na Contagem da População 1996.

Assim, alternativas para execução da captura de dados foram consideradas, indo desde a parceria para a elaboração do projeto, passando por formas distintas de fornecimento dos equipamentos, manutenção, e operação, chegando até a ser cogitada que a responsabilidade total pela execução dos servi-

ços ficasse a cargo do IBGE. A alternativa de aluguel de parte ou do total de equipamentos necessários, principalmente *scanners,* não foi descartada, visto que a experiência mostrava ser o parque necessário ao processamento de um censo, muito superior ao requerido para atividades contínuas da Instituição.

Por fim, optou-se pela aquisição dos equipamentos e dos serviços de manutenção dos mesmos, ficando a operacionalização da atividade de captura de dados a cargo do IBGE, que se utilizou da contratação de mão-de-obra temporária para a sua execução.

Após essas definições, forem determinados o número de Centros de Captura de Dados, suas localizações, os projetos físicos necessários e elaborados os respectivos editais para as contratações.

O uso da tecnologia de digitalização de documentos e reconhecimento ótico de marcas e caracteres trouxe vantagens, entre as quais pode-se destacar: a preservação do conteúdo dos questionários e Folhas de Coleta; a redução do fluxo de papel; a redução do tempo de captura de dados; o aumento da qualidade do processo de captura de dados; e a possibilidade de posterior visualização das imagens do material coletado.

3.2.1 Adaptação dos instrumentos de coleta à tecnologia

O uso, na captura dos dados do Censo, da leitura de documento por *scanner* – documento digitalizado – com posterior reconhecimento de caracteres através de *softwares* ICR – *Intelligent Caracter Recognition* ou OCR – *Optical Caracter Recognition* – em substituição aos processos de digitação e leitura de marcas óticas utilizadas em operações anteriores, demandou uma mudança significativa na confecção, manuseio, preenchimento, transporte e armazenamento dos instrumentos de coleta utilizados para registro dos dados.

3.2.1.1 Confecção dos instrumentos de coleta para registro dos dados

Com a adoção da tecnologia, uma nova forma de desenho foi aplicada a esses instrumentos de coleta. Seus campos foram milimetricamente delimitados e separados por um contraste de cores em formas retangulares, restringindo o preenchimento de cada caractere ao campo determinado.

Desenhos de códigos de barras foram introduzidos para identificação dos instrumentos que teriam seus dados capturados para o meio magnético – escaneados e armazenados – pois, desta forma, os diversos formulários – folhas diferentes de um mesmo instrumento – poderiam ser misturados durante o processo de escaneamento. Pequenas marcas em forma de quadrados pretos também foram introduzidas nas proximidades de suas margens, como pontos de balanceamento para permitir que os mesmos fossem posicionados durante o processo de digitalização. Tais instrumentos também receberam instruções e desenhos, com exemplos sobre como preencher corretamente os campos.

Na impressão, fez-se a utilização de cores invisíveis para todo o instrumento, com fundo branco, para as áreas de preenchimento; azul pastel e verde, normalmente, são cores invisíveis – *drop-out colors.* Ignoradas pelos *scanners* através de filtros, não permitem que sejam carregadas para as imagens armaze-

nadas em meio magnético – no caso do Censo Demográfico 2000, foi utilizada a cor azul pastel para todos os instrumentos de coleta digitalizados.

A gramatura do papel também foi observada na confecção desses instrumentos, pois os preenchimentos não poderiam ser identificados – deixar marcas – no seu verso, sendo para isso utilizado o papel *off set* de 90 gramas.

3.2.1.2 Preenchimento dos instrumentos de coleta

Assim como no desenho e na impressão, o cuidado com o preenchimento dos instrumentos é de fundamental importância na obtenção de um maior índice de reconhecimento de caracteres. Por esse motivo, várias adaptações foram realizadas nos manuais de instrução e nos treinamentos dos Recenseadores e Supervisores, responsáveis pelo preenchimento, com o intuito de que este fosse feito dentro dos padrões predeterminados, como, por exemplo: evitar rasuras; respeitar os limites dos campos; cuidar da caligrafia; e obedecer as regras de grafia – por exemplo, grafar o algarismo sete com um corte, de modo a evitar a possibilidade de identificação com o algarismo um.

No Censo 2000, optou-se pela utilização do lápis para registro dos dados, pois o uso da caneta só permitia correções com tinta branca, o que traria problema ao processo de escaneamento – liberação de fragmentos no equipamento. O grafite indicado pelos especialistas foi o de número dois, porque permite um registro claro da informação, sem danificar o papel.

Com a adoção do lápis, foi necessário a introdução da borracha como instrumento de correção. Sua escolha também passou por critérios que consideraram tamanho, maciez e a não liberação de resíduos durante o uso.

3.2.1.3 Manuseio, armazenamento e transporte dos instrumentos de coleta

Com a nova tecnologia e os novos instrumentos, o seu manuseio, armazenamento e transporte também passaram por adaptações. Para o manuseio e transporte durante o trabalho de coleta, foram confeccionadas pastas e pranchetas de material resistente.

O cuidado na confecção da pasta teve como objetivo dar condições para que os documentos não sofressem alterações em sua forma original. Já com a prancheta, buscou-se evitar que os Recenseadores utilizassem outro tipo de apoio, com exceção de uma mesa, para o devido registro da informação coletada.

Para o armazenamento e transporte dos questionários e Folhas de Coleta para os Centros de Captura de Dados, também com o objetivo de preservação das suas características originais, foram confeccionadas pastas de papelão grosso com as dimensões exatas para o acondicionamento dos documentos, com vistas a mantê-los protegidos após o fechamento da coleta do setor – cada pasta foi dimensionada para armazenar em média um setor. Cumpre salientar que, para o transporte, essas pastas foram também armazenadas em caixas maiores, evitando, mais ainda, danos ao material.

Essas adaptações relativas ao manuseio, armazenamento e transporte tiveram como preocupação o aumento da produtividade no processo de captura dos documentos, ou seja, fazendo com que a taxa de erros de reconhecimento fosse reduzida. Essa atividade de reengenharia dos processos em questão foi motivada pela preocupação de eliminar, ou reduzir ao máximo, as possibilida-

des de ocorrência de instrumentos preenchidos de forma indevida e, principalmente, de evitar a ocorrência de documentos amassados, rasgados, ou manuseados em excesso.

3.3 Sistemas de controle e apuração

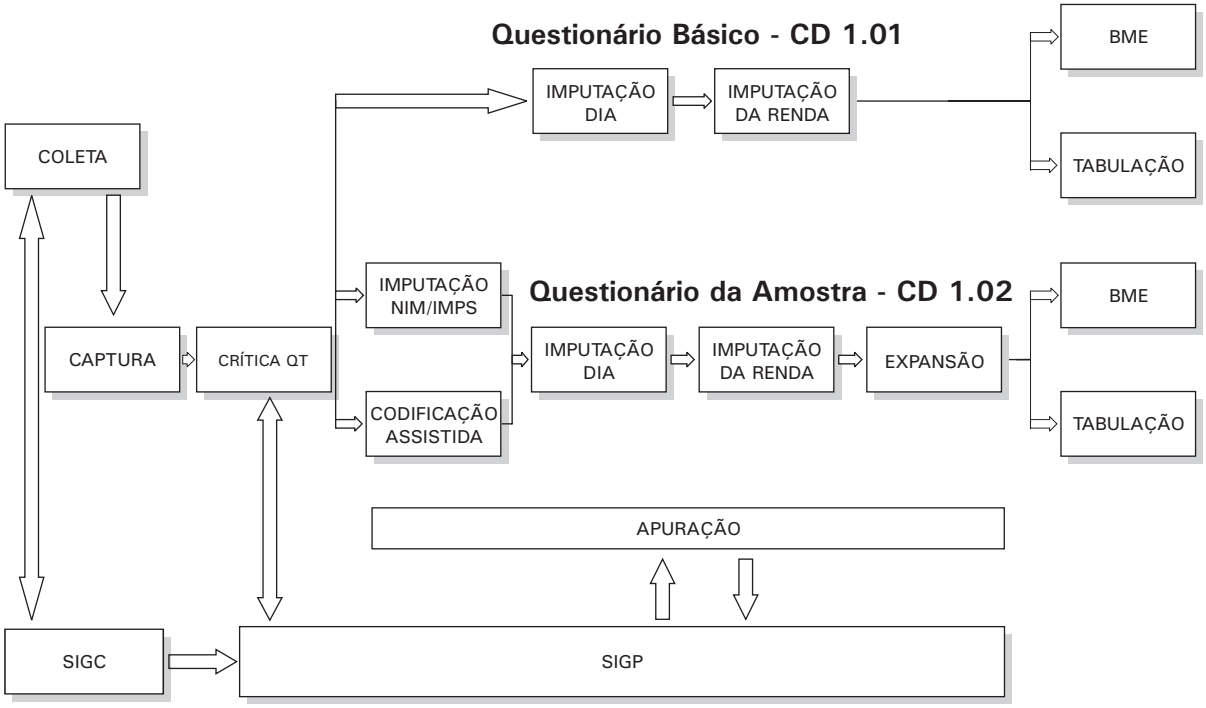
O Censo Demográfico 2000 envolveu o trabalho direto de cerca de 200 mil pessoas, entre contratados temporários e servidores efetivos do IBGE. Mais de 169 milhões de pessoas foram pesquisadas em aproximadamente 42 milhões de domicílios, nos 5 507 municípios, gerando 49 milhões de questionários preenchidos. Os questionários foram remetidos aos cinco Centros de Captura de Dados, onde tiveram seus dados convertidos para meio magnético. Posteriormente, foram trabalhados nas etapas de crítica, codificação, imputação, expansão, tabulação e análise, para então serem postos à disposição da sociedade.

A fim de manter o controle, assegurar a qualidade nos processos de trabalho e garantir o adequado gerenciamento de uma operação que envolveu expressivas quantidades de pessoas, material e informações, abrangendo em determinados períodos todo o território nacional, fez-se necessária a implementação de diversos sistemas de controle e apuração. Entre eles, destacam-se: o Sistema de Indicadores Gerenciais da Coleta – SIGC – o Sistema de Indicadores Gerenciais da Produção - SIGP - os vários sistemas utilizados nas etapas de captura de dados, crítica, codificação, imputação, expansão dos dados e, finalmente, na tabulação e disponibilização das informações em bases de dados. Todos esses sistemas estão descritos em detalhe nos capítulos e seções específicos de seus temas ao longo da publicação, e esta seção apresenta cada um deles sob o enfoque da logística de suas aplicações, nas diversas etapas da operação censitária.

Os sistemas SIGC e SIGP, em razão da função primordial de controle das grandes etapas operacionais – coleta e processamento dos questionários e Folhas de Coleta – tomaram por base as estruturas de organização do trabalho, segundo os recortes espaciais e as estruturas gerenciais, que significou a subdivisão das 27 Unidades da Federação e seus 5 507 municípios em 848 subáreas de coordenação e 6 828 postos de coleta – grupos de setores censitários, cada um deles atribuído a um Agente Censitário Municipal – ACM – conforme já descrito na seção 3.1.4, Estrutura Gerencial. Ambos os sistemas tiveram como unidade de controle o setor censitário e respeitaram a divisão político-administrativa vigente em 1º de janeiro de 2000.

Mil postos contaram com microcomputadores, impressoras e linhas de comunicação, e os ACMs neles sediados incluíam diretamente no SIGC as informações sobre o andamento da coleta e os dados de fechamento de cada setor. Já os ACMs dos postos não-informatizados repassavam os dados aos respectivos Coordenadores de Subárea, que os introduziam no sistema a partir de suas unidades de lotação – agências ou postos informatizados. A figura a seguir apresenta as grandes etapas de trabalho do Censo, que contaram com os sistemas de controle e apuração.

Figura 3.3 - Etapas de coleta e processamento dos questionários do Censo 2000



3.3.1 Sistema de Indicadores Gerenciais da Coleta

O Sistema de Indicadores Gerenciais da Coleta - SIGC foi implementado por meio das tecnologias Lotus Notes*, Web e Crystal Report*, para atender o gerenciamento da coleta do Censo Demográfico 2000. O sistema era acessado na sede do IBGE e das Unidades Regionais, bem como em qualquer de seus postos informatizados, segundo uma hierarquia de acessos, através de senhas de segurança.

O SIGC foi implementado com o propósito de atender aos seguintes objetivos:

- permitir a todos os níveis da estrutura censitária e à direção do IBGE acesso rápido e *online* a relatórios que demonstrassem a situação do andamento da coleta dos dados;

- oferecer às coordenações técnicas, ao longo do trabalho, indicadores de avaliação da coleta, que permitissem a atuação rápida em caso de identificação de alguma distorção;

- fornecer os Resultados Preliminares do Censo, tão logo fosse encerrada a coleta dos dados;

- promover a integração com os sistemas administrativos que iriam apoiar a operação censitária, especialmente com o de pagamento dos Recenseadores; e

- criar uma base de dados de referência para os trabalhos nos Centros de Captura de Dados.

No que se refere à integração, o sistema possibilitou a aproximação das unidades do IBGE em todo o País. Durante toda a operação de coleta, os equipamentos foram interligados com o auxílio da Empresa Brasileira de Teleco-

municações - Embratel, que disponibilizou um provedor para a operação – *IP Office Dial* – mantendo seguras todas as informações que trafegavam pela linha de comunicação.

Em relação ao pagamento dos Recenseadores, o sistema permitia conhecer exatamente a produção de cada um deles, sendo possível integrá-la, de forma automatizada, com à execução da folha de pagamento, que passou a ser emitida, em média, após cinco dias do término da coleta do setor.

A manutenção do sistema foi também facilitada, uma vez que tudo foi feito através da *Web*. O programa *Workflow* permitia solucionar qualquer tipo de problema via Internet, com o máximo de segurança.

Para auxiliar na apresentação das informações gerencias, foram utilizados recursos do programa *Crystal Report*, ferramenta útil para o reconhecimento de dados, análise de conteúdo e geração de gráficos, relatórios e tabelas.

Uma descrição mais detalhada do sistema, bem como do uso de suas informações, pode ser encontrada no Capítulo 12.

3.3.2 Sistema de Indicadores Gerenciais da Produção

O Sistema de Indicadores Gerenciais da Produção – SIGP foi criado, principalmente, para fornecer informações sobre o andamento dos trabalhos de apuração das informações dos Questionários Básico e da Amostra, ou seja, da codificação e da crítica e expansão dos dados. Servia também para controlar a remessa do material coletado feita pelas Unidades Regionais para os Centros de Captura de Dados.

Quando da remessa de material – Questionários Básico, da Amostra e Folhas de Coleta – eram preenchidas guias de remessa eletrônica através do SIGP/Intranet. Assim, automaticamente, os CCDs tomavam conhecimento do envio através do sistema, permitindo planejar o recebimento e a guarda do material remetido

No que se refere à apuração, foi definido um banco de controle, com informações sobre a formação dos lotes de trabalho – número de setores, de questionários, de pessoas, etc. – bem como sobre datas e situações de processamento – liberado para execução, execução com problemas, execução com sucesso, etc.

O banco de dados foi implementado com tecnologia *Oracle* e era atualizado automaticamente durante o processo de apuração. As informações eram repassadas, periodicamente e de forma automática, para a visão SIGP/*Web* e disponibilizadas através de relatórios gerenciais, implementados com tecnologia *Crystal Report* via *Web*/Intranet.

3.3.3 Sistemas de captura de dados

As tecnologias utilizadas para a captura de dados do Censo 2000 foram a digitalização por *scanner* dos questionários e Folhas de Coleta, bem como o reconhecimento ótico de código de barras, marcas e caracteres, que mostraram ser as melhores alternativas para o processamento de grandes volumes de documentos.

A captura dos dados foi feita nos CCDs, de forma simultânea, para os Questionários Básico e da Amostra. O quadro abaixo mostra as Unidades da Federação, cujo material coletado foi processado em cada Centro de Captura de Dados.

Quadro 3.2 - Centros de captura de dados e Unidades da Federação processadas

Centros de captura de dados	Unidades da Federação
Rio de Janeiro/RJ	Rondônia, Acre, Amazonas, Roraima, Pará, Amapá e Tocantins
Campinas/SP	São Paulo, Espírito Santo e Mato Grosso do Sul
Goiânia/GO	Minas Gerais, Goiás, Mato Grosso e Distrito Federal
Curitiba/PR	Paraná, Santa Catarina, Rio Grande do Sul e Rio de Janeiro
Campina Grande/PB	Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe e Bahia

Durante a operação de captura, que durou aproximadamente 100 dias úteis, foram escaneados e processados aproximadamente 200 milhões de páginas do tipo A3. Cada CCD foi configurado conforme suas necessidades de processamento, sendo que a menor unidade de operação era o módulo de reconhecimento, composto por *scanners*, equipamentos servidores de imagem e dados, micros de verificação e crítica, além de processadores de reconhecimento de caracteres.

Os módulos foram instalados nos CCDs, conforme as quantidades descritas na tabela abaixo.

Tabela 3.4 - Configuração dos centros de captura de dados

Centros de captura de dados	Módulos	*Scanner*	Equipamento servidores	Micros verificação	Micros crítica	Equipamento reconhecimento
Rio de Janeiro	1	2	3	18	6	8
Goiás	2	5	5	36	12	16
Paraíba	4	8	9	72	24	32
Paraná	4	8	9	72	24	32
São Paulo	4	8	9	72	24	32

Fonte: IBGE, Diretoria de Informática.

3.3.4 Imputação NIM/IMPS

O *New Imputation Methodology* – NIM foi utilizado para crítica e imputação automática de um conjunto de variáveis do Questionário da Amostra, referentes às pessoas residentes em domicílios com até oito moradores. O processamento foi executado em plataforma *Windows*, através de lotes formados por estratos de domicílios, de acordo com o número de pessoas. Nos domicílios com mais de oito pessoas, as correções foram feitas manualmente, através de uma aplicação desenvolvida pelo *Integrated Microcomputer Processing System* – IMPS, *software* gratuito desenvolvido pelo U.S. Census Bureau. Além disso, se em algum dos lotes processados pelo NIM surgissem domicílios cujas variáveis não estivessem corrigidas, os registros correspondentes eram encaminhados para o processamento pela aplicação IMPS.

O NIM foi cedido pelo Statistics Canada e tinha sido utilizado na Contagem de População 1996 daquele País. Consiste de um conjunto de programas e rotinas em linguagem de programação C para uma aplicação específica de cinco variáveis. Não é um sistema genérico, o que significa que, para cada aplicação em um conjunto de variáveis diferente, é necessário efetuar adaptações nos programas e rotinas originais em C. No entanto, toda a parte do *software* referente à metodologia de imputação pode ser aproveitada. Isto significa que, para cada nova aplicação, somente é necessário fazer as alterações nas partes do sistema de entrada – leitura – armazenamento e saída dos dados, sem necessidade de alterar os algoritmos de imputação.

Inicialmente, o sistema foi adaptado de forma a ser executado no compilador C disponível, para um conjunto inicial de variáveis do Censo Demográfico 2000. Posteriormente, foram sendo feitas as adaptações para as aplicações de crítica da Segunda Prova Piloto, do Censo Experimental e do Censo 2000. Após as adaptações necessárias no código C, foi gerado um arquivo executável, instalado no ambiente de produção *Windows NT*, para processamento da crítica.

3.3.5 Codificação

A fase de Codificação do Censo Demográfico 2000 foi realizada com a implementação de dois sistemas: de Aplicação de Códigos e de Verificação de Códigos. Nesses sistemas, as unidades de trabalho eram os mesmos lotes constituídos para a crítica dos dados, sendo tratadas as descrições registradas no Questionário da Amostra, coletadas em aberto, para os quesitos relativos aos temas Religião, Migração, Curso Concluído, Ocupação e Atividade.

Os sistemas eram operados em microcomputador e permitiam o acompanhamento do processo de produção, apresentando informações por lote e tema trabalhados, ao mesmo tempo que incorporavam mecanismos automáticos de atualização dos dados.

Informações completas sobre a Aplicação e Verificação de Códigos podem ser obtidas na Seção 12.5.1.

Aplicação de códigos

O modelo utilizado na aplicação de códigos foi uma combinação entre os processos manual e automático. Representou algo como uma pós-codificação, em que os códigos para cada quesito só eram atribuídos após o conhecimento de todas as descrições encontradas nos questionários de um determinado lote de trabalho. O sistema apresentava as seguintes etapas: extração de textos, correção de caracteres, agrupamento por códigos, confirmação de códigos e aplicação aos registros individuais. Suas características principais são as seguintes:

- as descrições encontradas nos questionários eram armazenadas individualmente para cada pessoa;

- essas descrições eram transferidas para uma base de dados em separado; nessa base, as informações semelhantes, para cada tema, foram unificadas e agrupadas segundo os critérios da aplicação de códigos – divisão em palavras, eliminação de preposições, artigos, plurais e gênero, substituição de letras de mesmo som e uso de sinônimos;

- a aplicação dos códigos foi feita com base nas descrições por lote; todos os registros com as mesmas características tiveram os códigos aplicados de uma única vez;

- a aplicação de um código era, inicialmente, confirmada pelo codificador do tema, mediante a validação do código proposto pelo sistema; após ser atingida, para uma mesma situação de confirmação, uma determinada freqüência e, após a validação de um orientador, a aplicação do código passava a ser feita de forma automática, sem a necessidade de confirmação; e

- terminada a aplicação de códigos do arquivo de descrições do lote, o arquivo original – lote de codificação – em que se incluem os registros individualizados, e o arquivo com as descrições já codificadas passaram por um batimento, visando atribuir os códigos aos registros de cada pessoa.

Verificação de códigos

A verificação de códigos teve como objetivo contribuir para a melhoria da qualidade do trabalho da etapa Aplicação de Códigos. Operacionalmente, a tarefa consistiu em fazer, numa amostra de registros, uma segunda atribuição de códigos, para todos os lotes de cada quesito, visando identificar as situações de divergência entre os códigos aplicados pelos dois operadores, codificador e verificador.

As principais características do sistema de Verificação de Códigos são listadas abaixo:

- liberado um lote de codificação, era constituído o correspondente lote de verificação, que continha somente os registros das pessoas em que não houve a aplicação automática de códigos, ou seja, onde aconteceu a intervenção do codificador – na etapa de correção de caracteres e/ou confirmação de códigos;

- no lote de verificação, era selecionada uma amostra de registros onde o verificador, utilizando a mesma rotina de trabalho do codificador, com exceção dos mecanismos de correção e confirmação automáticos, fazia a nova aplicação de códigos; e

- eram identificadas as situações de divergência entre os códigos aplicados pelo codificador e verificador, e calculada sua proporção na amostra;

Para todos os lotes cuja proporção de códigos divergentes superasse um valor preestabelecido, fazia-se uma investigação com o propósito de esclarecer o que acarretou as divergências entre os códigos aplicados, ou seja, se houve erro do codificador e/ou do verificador. Dependendo do resultado da investigação, eram tomadas providências – alertas, retreinamento de pessoal, etc. – junto aos codificadores, orientadores e verificadores, tendo em vista melhorar a qualidade do trabalho.

3.3.6 Imputação DIA

OSistema de Detección e Imputación Automática de Errores para Dados Cualitativos – DIA desenvolvido pelo Instituto Nacional de Estatística – INE – da Espanha, tem como base a metodologia de Fellegi e Holt (1976) – de imputação probabilística – que trata dados qualitativos inconsistentes ou ausentes, aliando também a possibilidade de serem feitas imputações determinísticas.

Esse sistema foi operado em ambiente centralizado, tipo *mainframe* IBM, estando constituído por dois subsistemas: Subsistema de Especificação e Subsistema de Tratamento. Sua implementação garante que as correções de erros em um registro – conjunto de variáveis relativas a uma pessoa ou domicílio - possam ser realizadas de uma única vez, de modo que sejam atendidas todas as funções de crítica estabelecidas, evitando-se assim, os ciclos habituais de crítica, que podem ocorrer quando da utilização de outros programas de imputação.

O Censo Demográfico 2000 utilizou o sistema DIA para a crítica e imputação das informações do Conjunto Universo – descrito na Seção 12.3.1 Formação do Conjunto Universo – e, também, para a crítica intra-registros das informações pesquisadas no Questionário da Amostra.

Para a implementação do processo de crítica/imputação, foram definidos lotes de trabalho, através de critérios de formação estabelecidos separadamente para os lotes do Conjunto Universo e dos dados do Questionário da Amostra.

O DIA era implementado através de aplicativos específicos, que realizavam a detecção e imputação dos erros para as informações investigadas por um determinado conjunto de quesitos. Esses quesitos podiam pertencer a um único bloco de investigação do questionário, ou a mais de um. Assim, foram definidos aplicativos cujas funções de crítica investigavam, em separado, por exemplo, as variáveis dos temas Fecundidade, Migração, Trabalho, etc.

Uma descrição completa do uso do DIA no processo de crítica e imputação dos dados do Censo Demográfico 2000 pode ser encontrada nas seções 12.3 e 12.5 do Capítulo Apuração dos Dados.

3.3.7 Imputação da renda

A imputação da renda foi executada pelo *software* S-PLUS em ambiente operacional *Windows 98*. O procedimento de imputação implementado foi baseado em árvores de regressão e está descrito nas Seções 12.3.2.7 e 12.5.2.4. A unidade de produção dessa etapa da apuração, consistiu do mesmo lote de questionários usado na imputação realizada pelo DIA. Como os arquivos correspondentes aos lotes se localizavam no ambiente operacional OS/390 – *mainframe* IBM – foi desenvolvido, para conexão entre esses dois ambientes, através do *software Statistical Analysis Sistem* – SAS – uma rotina computacional para automatizar todo o processo da imputação, constituído das seguintes etapas, que foram executadas em cada lote.

Preparação do arquivo de entrada para o S-PLUS

Esta etapa era executada no *mainframe*. Era gerado um arquivo SAS com as variáveis de rendimento e um conjunto de variáveis explicativas dos rendimentos dos moradores, além da identificação de cada pessoa no domicílio. Esse arquivo, que não possuía registros para as pessoas menores de dez anos, era transferido, via *PROC DOWNLOAD* do SAS, para o ambiente *Windows 98*.

Imputação da renda

Nesta etapa, o programa SAS ativava o *software* S-PLUS para processar o arquivo preparado na etapa anterior.

Transferência dos resultados para o ambiente OS/390

O S-PLUS produzia como resultado um conjunto de relatórios e um diário de ocorrências de execução, que eram transferidos ao *mainframe* por FTP – *File Transfer Protocol*. Também era gerado um arquivo SAS, semelhante ao recebido como entrada, com as variáveis de rendimento imputadas. Esse arquivo era, então, transferido ao mainframe via *PROC UPLOAD* do SAS.

Atualização do lote no ambiente OS/390

Esta etapa era executada no *mainframe*. O programa SAS preparava um arquivo com comandos da *Job Control Language* – JCL – do OS/390, e o submetia a execução. Este *job* ativava um programa que recebia como entrada o arquivo do lote original e o arquivo SAS gerado pelo S-PLUS, e promovia a atualização das variáveis imputadas.

3.3.8 Expansão da amostra

A expansão da amostra do Censo Demográfico 2000 foi realizada por um conjunto de procedimentos envolvendo microcomputador e *maiframe*, através de programas especialmente desenvolvidos. A parte processada em microcomputador refere-se à definição das áreas de ponderação usadas na calibração dos pesos amostrais. Cada área consistiu de uma unidade geográfica formada por um agrupamento de setores censitários, utilizada para a aplicação dos procedimentos de calibração das estimativas, obedecendo às informações conhecidas para a população como um todo – Conjunto Universo. Detalhes sobre a formação das áreas de ponderação podem ser vistos na Seção 12.5.3 Expansão da Amostra.

Foram definidas para todo o Brasil 9 336 áreas de ponderação e, tal como nos censos anteriores, a metodologia de expansão da amostra foi aplicada separadamente, para cada uma delas. Arquivos criados no *mainframe* com as áreas definidas em cada uma das Unidades da Federação, serviram para controlar a geração dos pesos amostrais calibrados para os domicílios e pessoas.

Baseado nesses arquivos, foi desenvolvido um aplicativo em Oracle, executado em microcomputador, para controle do trabalho da expansão. Esse aplicativo gerenciava uma tabela onde se podia saber a situação de cada uma das áreas de ponderação em relação ao trabalho, ou seja, se estava liberada para executar a ponderação, se a ponderação já havia sido executada, se os pesos gerados eram satisfatórios, se era necessário novo processamento etc.

Foram ainda desenvolvidos dois procedimentos para execução das tarefas no *mainframe.* O primeiro rodava os programas em SAS para a criação dos arquivos preliminares, com as variáveis auxiliares do método de expansão utilizado. Esse procedimento era realizado separadamente para cada Unidade da Federação, assim que estivessem definidas todas as suas áreas de ponderação. Nessa oportunidade, eram gerados os dois arquivos básicos do sistema, contendo o valor das variáveis auxiliares para cada um das unidades domiciliares do universo e da amostra.

O segundo procedimento executava, por meio de programas desenvolvidos em SAS, o cálculo dos pesos amostrais calibrados para uma dada área de ponderação. Como saída desse procedimento, era gerado um arquivo com os pesos calibrados a serem atribuídos a cada um dos domicílios da amostra e seus respectivos moradores.

3.3.9 Tabulação Pégasus

A apresentação dos dados em tabelas é uma das maneiras de oferecer ao usuário final os resultados de uma pesquisa, permitindo a consulta às variáveis de modo fácil e rápido.

O avanço tecnológico introduzido na tabulação do Censo Demográfico 2000 consistiu na confecção de tabelas com o uso do sistema genérico de tabulação Pégasus, desenvolvido no IBGE em linguagem *Visual Basic* e banco de dados FACIL – TSGDB. Antes disso, as tabulações eram realizadas através de programas "ad-hoc", desenvolvidos em linguagens Cobol e PL/I ou outras específicas para tabulação – Atlas e Prometeu – também desenvolvidas no IBGE. O uso desses sistemas quase sempre levava ao não-cumprimento dos prazos estabelecidos, uma vez que, para qualquer erro detectado, demandava-se tempo elevado para correção.

A solução implementada no Pégasus apresenta vantagens que tornam o trabalho mais ágil, rápido e com menor possibilidade de erro. Entre todas as vantagens, quatro foram de fundamental importância para seu uso no Censo 2000:

- facilidade para desenhar tabelas;

- simplicidade para montagem dos algoritmos para geração de variáveis derivadas – por grupamento, recodificação, cálculo de valores, etc.;

- redução considerável dos tempos de desenvolvimento e correção; e

- Interface com programa de planilha Excel, facilitando o trabalho de análise e conferência dos valores apresentados.

A arquitetura do Pegasus é cliente-servidor, centralizando seu banco de dados em um servidor RISC, sob gerenciador OpenBASE X-7.3.

3.3.10 Banco Multidimensional de Estatísticas

A preservação do sigilo das informações individualizadas, sejam elas de pessoas, famílias ou empresas, a que está submetido o IBGE por força de lei, cria restrições ao acesso público à sua base de dados, onde estão armazenados os microdados – os dados coletados nos questionários.

Para conciliar a preservação do sigilo com as demandas externas por informações mais desagregadas e ainda facilitar o acesso, a recuperação e o trabalho dos dados, foi criado o Banco Multidimensional de Estatísticas – BME. Nele, cada questionário constitui um ou mais registros de dados, sendo, portanto, a informação mais desagregada possível em uma pesquisa estatística. O resultado de uma consulta, no entanto, é uma tabela agregada que preserva o sigilo legal da informação. Desta forma, em comparação com censos anteriores, o BME introduziu uma revisão conceitual na organização da base de dados censitários, com a incorporação de métodos seguros para o trabalho de infor-

mações desagregadas e um serviço ativo de metainformação, que se constitui na descrição das variáveis e dos conceitos utilizados na investigação, entre outras informações.

Todas as variáveis pesquisadas no Censo Demográfico 2000 foram disponibilizadas no BME, agrupadas por temas e subtemas.

3.4 Centros de Captura de Dados

Tão logo foi decidida a utilização da tecnologia de reconhecimento ótico com uso de imagens digitalizadas para o processamento dos documentos do Censo Demográfico 2000, foram iniciados estudos com vistas a identificar os aspectos logísticos relevantes para os processos de trabalho inerentes à tecnologia que era, então, nova para o IBGE.

Em fins de abril de 1998, foi feita uma visita ao U.S.Bureau of the Census, em Washington, nos Estados Unidos, para se conhecerem detalhes da solução recentemente escolhida. Entrevistaram-se os gerentes das Áreas de Pesquisas e de Informática – desenvolvimento e produção – e pôde-se conhecer aplicativos de acompanhamento do Projeto Censo e de codificação automática.

Foram feitas reuniões com as empresas contratadas para fornecer as soluções de *hardware* e *software* para o Censo e também para a instalação dos Centros de Captura de Dados. Finalmente, foi visitado o Centro de Produção em Jeffersonville, onde estava instalada a solução utilizada no Censo Experimental americano. Ali, foi possível conhecer alguns aspectos operacionais da apuração de pesquisa com a utilização da tecnologia de reconhecimento ótico de caracteres. Um deles foi a montagem da operação em módulos, que permitia segmentar os recursos físicos e humanos, a fim de obter melhor desempenho da produção, evitando, assim, que eventuais problemas ocorridos causassem a paralisação da operação inteira.

Essa visita permitiu que fossem atualizados os estudos realizados até então, e ajudou a definir algumas questões importantes sobre a forma de apuração do Censo 2000. As primeiras conclusões indicaram parâmetros e sugeriram procedimentos que nortearam as ações adotadas para a montagem das instalações físicas onde foram processados os instrumentos de coleta, os Centros de Captura de Dados – CCDs. Os principais parâmetros e procedimentos foram:

- a quantidade de centros deveria ficar restrita a quatro ou cinco unidades, de forma a otimizar a capacidade de produção de cada uma delas, garantir a uniformidade de procedimentos, reduzir gastos administrativos com infra-estrutura e equipamentos, otimizar a alocação de pessoal técnico especializado e seus deslocamentos, bem como assegurar supervisão efetiva por parte dos órgãos centrais do IBGE;

- os centros deveriam estar situados, preferencialmente, em cidades de porte médio a grande – com cerca de 500 mil habitantes – que fossem centros universitários, oferecessem baixos custos de locação e mão-de-obra de qualidade, além de dispor de facilidades de acesso e comunicação e de transporte coletivo noturno;

- um dos centros deveria ser localizado na cidade do Rio de Janeiro, preferencialmente no complexo de Parada de Lucas, onde já funcionavam unidades do IBGE, sendo este o primeiro CCD a ser montado, a fim de servir não apenas para a apuração dos questionários do Censo Experimental, mas ainda como modelo para a instalação e funcionamento dos demais. O CCD do Rio de Janeiro, vinculado à Diretoria de Informática, deveria ser também o responsável pelo processamento do material da Região Norte do País, e, por estar próximo dos órgãos centrais do IBGE, servir como laboratório para incorporação da tecnologia e, posteriormente, permanecer em funcionamento para atender a outros trabalhos da Instituição;

- os CCDs deveriam ser dimensionados de forma a cumprir todas as suas etapas de trabalho em cerca de 100 dias úteis;

- os centros deveriam ser planejados para operar em até três turnos de trabalho, com carga horária mínima de 14 horas diárias;

- o IBGE deveria adquirir os equipamentos, *softwares* e aplicativos necessários;

- para fins de instrução dos processos licitatórios, com vistas à contratação de serviços de desenvolvimento de *softwares* e aplicativos, deveriam ser definidos os níveis de confiança exigidos, estabelecendo-se as taxas mínimas de reconhecimento e as taxas máximas de erros que a solução deveria proporcionar para os tipos de campo dos questionários a serem lidos;

- os centros deveriam ser unidades autônomas, vinculadas diretamente aos órgãos centrais da Instituição;

- os gerentes - Coordenador Geral e coordenadores setoriais – dos centros deveriam ser servidores efetivos do IBGE, escolhidos em processo seletivo interno;

- as empresas contratadas deveriam manter, em cada CCD, pessoal técnico especializado para gerenciamento, suporte e supervisão dos respectivos sistemas; e

- a escolha dos locais – instalações prediais – a locação dos imóveis, a infra-estrutura de refrigeração, iluminação e energia elétrica seriam de responsabilidade do IBGE. Aos fornecedores caberiam todas as demais instalações para o funcionamento dos sistemas.

A Comissão de Planejamento e Organização Geral do Censo 2000 apreciou os resultados dos estudos realizados, discutiu os parâmetros e procedimentos propostos e aprovou a implantação do projeto, com cinco Centros de Captura de Dados, quais sejam: CCD Campina Grande – PB, CCD Campinas – SP, CCD Curitiba – PR, CCD Goiânia – GO e CCD Rio de Janeiro – RJ.

A Resolução do Presidente, R.PR nº 0001, de 1º de março de 2000, constituiu os cinco Centros de Captura de Dados do Censo 2000, subordinando-os administrativamente à Coordenação de Acompanhamento e Controle Operacional dos Censos, com orientação técnico-normativa das Diretorias de Informáti-

ca e de Pesquisas, segundo os assuntos inerentes a cada uma delas, cabendo à Diretoria-Executiva a emissão dos instrumentos normativos necessários à sistematização do funcionamento dos centros.

3.4.1 Instalação Física dos Centros

Os Centros de Captura de Dados de Campina Grande, Goiânia, Campinas e Curitiba foram montados para funcionar somente durante a execução do Censo Demográfico 2000. Para a instalação desses centros, algumas características foram cuidadosamente observadas, levando-se em conta os seguintes aspectos:

- a localização – facilidade de descarregar o material de coleta a qualquer hora; ter meio de transporte coletivo, em horário compatível com o funcionamento do centro; facilidade de estacionamento de veículos; proximidade de restaurantes para uso do pessoal; não sujeito a enchentes no local ou nas imediações; afastado de imóveis que manipulassem produtos inflamáveis;

- o imóvel, inclusive documentação – o imóvel deveria estar livre de qualquer ônus – impostos, taxas, etc; o proprietário, pessoa física ou jurídica, deveria ter registro no Sistema de Cadastro de Fornecedores – SICAF; e

- as instalações e as condições físicas – existência de energia elétrica ou possibilidade de sua instalação; regularidade de abastecimento de água; viabilidade de linhas de telefonia; condições hidráulicas adequadas; possibilidade de aclimatação; bom estado dos sanitários e probabilidade de ampliação, se necessário; instalações contra incêndio; verificação de colunas; o imóvel deveria ter um único piso.

Após a seleção dos imóveis, a equipe de engenharia do IBGE preparou, em conjunto com a Diretoria de Informática, projetos de adaptação física às atividades típicas dos CCDs. As plantas das instalações físicas de cada um dos centros estão apresentadas no CD-ROM anexo a esta publicação.

As licitações e as obras necessárias sofreram atrasos e foram as principais razões do adiamento do início dos trabalhos dos CCDs. Paralelamente às obras de adaptação, foram feitas as licitações para aquisição de mobiliário e outros equipamentos de infra-estrutura, tais como: mesas, cadeiras, arquivos, prateleiras, aparelhos de ar condicionado, etc. Foram também compradas dez máquinas tira-grampos especialmente desenvolvidas e construídas para a operação do Censo. Os recursos de informática – microcomputadores, servidores, impressoras, *switchs*, etc. – foram objeto de licitação específica que incluía também suporte *on-site* durante todo o período de apuração do Censo.

3.4.2 Seleção interna dos coordenadores dos centros

A decisão de implantar os Centros de Captura de Dados como unidades autônomas implicou a criação da função de Coordenador Geral de CCD. Foi realizado, então, um processo seletivo interno para selecionar servidores com perfil adequado às funções de coordenação dos centros. Esta Seção identifica os instrumentos usados e discrimina os critérios de avaliação para a determinação dos candidatos com potencial para ocupar as vagas.

O servidor deveria ter capacidade gerencial para: administrar uma unidade técnico-operacional de grande porte; ser criativo, dinâmico e automotivado; ter iniciativa para diagnosticar problemas; ter habilidade e liderança para tratar com pessoas de diferentes níveis; ter competência emocional para atuar sob forte pressão de tempo e grande volume de trabalho; ter facilidade de comunicação; ter conhecimento de informática e noções de administração de recursos humanos, orçamentários e financeiros; e buscar soluções, tomar decisões e apresentar rapidez de assimilação dos processos de captura de dados que seriam utilizados no Censo 2000.

No processo seletivo interno, foi feita a avaliação das qualificações dos 123 servidores/candidatos inscritos, ao longo das seguintes etapas:

- Currículo Dirigido – elaborou-se um currículo dirigido, visando à investigação de requisitos básicos para o perfil necessário para coordenador do centro: formação de 2º grau; conhecimento nas áreas de informática, administração de recursos humanos, orçamentários e financeiros; experiência profissional; expectativa em relação ao cargo; e preferências de lotação, já que existiam quatro centros com localização bastante diferenciada. Nesta etapa, os servidores/candidatos foram categorizados em *recomendados* e *não-recomendados*, e somente os 94 primeiros foram para a etapa seguinte do processo seletivo;

- Exercício Situacional – composto de uma prova única, aplicada em todas as unidades regionais que apresentassem candidatos, com questões inerentes às funções a serem exercidas no dia a dia de um centro, para avaliar o preparo dos candidatos quanto: à comunicação escrita; à tomada de decisões; à solução de problemas e à iniciativa frente às situações presentes. Esta etapa, com peso dois, classificou os 20 primeiros candidatos para a etapa seguinte;

- Dinâmica de Grupo – através deste exercício, pôde-se alcançar uma situação comparativa entre as pessoas e suas ações comportamentais, possibilitando o reconhecimento de seus valores individuais e grupais quanto: à criatividade; ao relacionamento; à liderança; ao trabalho sob pressão; à motivação; à comunicação oral; à autoconfiança: ao equilíbrio emocional. Os servidores/candidatos receberam avaliação com peso dois para compor a sua avaliação final;

- Entrevista Técnica – foi realizada por gerentes envolvidos na preparação do Censo, os quais, de posse de todas as informações relativas às etapas anteriores do processo seletivo interno, em entrevistas individuais, puderam conhecer melhor as características e as expectativas do servidor em relação ao cargo a ser exercido e avaliá-los. Para efeito de seleção esta etapa teve peso três;

- Entrevista Social – foi feito um levantamento da situação socioeconômica do servidor, visando avaliar sua futura adaptação na nova atribuição, bem como a necessidade da reestruturação de vida na localidade para a qual fosse designado com base nos seguintes dados: o histórico funcional dentro do IBGE, lotações anteriores, estrutura familiar atual e anterior, estrutura funcional atual e expectativas relacionadas ao trabalho;

- Exame Médico – foi feita a avaliação médica dos servidores/candidatos através da inspeção análise dos seus históricos médicos no IBGE; no ato da entrevista, exigiram-se exames, tais como: hemograma completo, EAS, raio X do tórax, glicemia, colesterol total LDL – HDL, triglicerídeos, uréia e creatinina. A avaliação médico-social caracterizava os servidores/candidatos em habilitados ou não-habilitados, compondo, assim, o dossiê para a classificação final; e

- Treinamento – esta etapa tinha o objetivo de qualificar e avaliar o servidor/candidato com base no trabalho a ser desenvolvido, constando do seu conteúdo programático, módulos gerenciais e técnicos. Participaram deste treinamento os dez primeiros classificados nas etapas anteriores. Esta etapa teve peso três, originando a média final do processo seletivo interno.

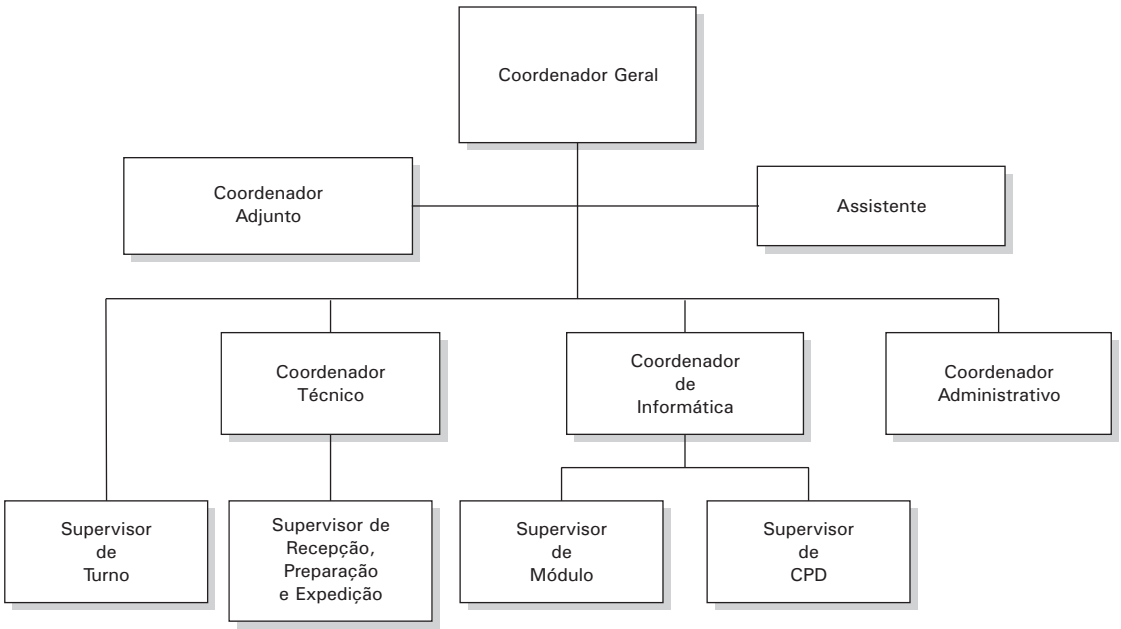
O referido processo selecionou os quatro primeiros candidatos para exercerem a função de Coordenador Geral dos Centros de Captura de Dados, com duração prevista para cerca de dois anos. A nomeação foi feita com um ano de antecedência em relação ao prazo previsto para início dos trabalhos de apuração, para que cada um deles pudesse participar de todo o processo de montagem do centro que iria administrar, visando, assim, à maior eficiência no seu gerenciamento.

O centro do Rio de Janeiro, diferentemente dos demais, por ser uma unidade permanente, a ser operada exclusivamente por servidores do IBGE para a absorção de conhecimentos e a tecnologia dos sistemas, teve o seu Coordenador Geral indicado pela Diretoria de Informática.

3.4.3 Organização e estrutura dos Centros de Captura de Dados

Com base na estruturação do CCD do Rio de Janeiro e com o acompanhamento de todos os Coordenadores Gerais, os centros foram montados com a seguinte estrutura de cargos:

Figura 3.4 – Organograma do Centro de Captura de Dados



- Coordenador Geral – responsável pela coordenação, direção, organização e planejamento geral no centro, além de representante junto à administração central;

- Coordenador Adjunto – substituto do Coordenador Geral e responsável por tarefas especiais por ele designadas;

- Coordenador Técnico – responsável pela apuração dos dados, garantindo sua qualidade e prazo de execução. Representante do CCD junto à Diretoria de Pesquisas;

- Coordenador Administrativo – responsável pelas atividades ligadas ao controle dos recursos humanos, materiais, patrimoniais e financeiros do CCD. Representava o CCD junto às áreas da Diretoria-Executiva;

- Coordenador de Informática – responsável por todas as atividades ligadas à infra-estrutura de informática instalada no CCD e pela segurança física dos dados apurados. Indicado para manter contato com os fornecedores de *hardware* e *software*. Representante do CCD junto à Diretoria de Informática;

- Supervisores de Turno – responsáveis pela abertura e fechamento dos trabalhos diários do CCD, pelo controle geral do andamento da operação, pela comunicação com os operadores e pelo relato dos problemas aos coordenadores;

- Supervisores da Recepção, Preparação e Expedição – responsáveis pela recepção, preparação e expedição de material; e

- Supervisores de Módulo – responsáveis pelos módulos operacionais; e

- Supervisores de Controle de Qualidade – responsáveis pela execução dos aplicativos de controle de qualidade do reconhecimento ótico de caracteres e pela emissão dos relatórios de acompanhamento do processo.

A ocupação desses cargos se deu primeiramente por indicação dos Coordenadores Gerais, utilizando servidores do quadro de pessoal do IBGE. Para os cargos de supervisão, onde não havia pessoal do quadro disponível, foram selecionados e treinados os Auxiliares Censitários melhores classificados no processo seletivo - ver Capítulo 3, Seção 3.1.5.1, item f.

Os CCDs foram organizados para atender às rotinas de trabalho de um processo de produção de captura de dados ao longo das seguintes etapas:

- Recepção e Expedição – recebimento das pastas vindas das Divisões Estaduais de Pesquisas e pela conferência de todo o material; registro no sistema; numeração das pastas; armazenamento e expedição do material trabalhado;

- Preparação – esta fase foi composta das seguintes etapas: pegar caixa no carrinho; retirar questionários da caixa; retirar grampos dos questionários; contar e abrir questionários; acondicionar o lote em bandeja; registrar o lote no sistema, e gerar a capa do lote. Qualquer anormalidade detectada durante a contagem do material – questionários com clips, rasgados, molhados, sem código de barras, etc. foi tratada de forma específica;

- Digitalização – nesta fase, fazia-se a digitalização dos formulários, separados por modelos, envolvendo quatro etapas: arrumar os questionários para escaneamento; passar os questionários pelo *scanner*; acompanhar as imagens digitalizadas na estação de digitação e enviar o lote para reconhecimento;

- Reconhecimento – nesta fase, totalmente automatizada, as imagens digitalizadas eram transformadas em arquivos de dados;

- Verificação – tinha o objetivo de tratar os dados que não tivessem atingido o índice de reconhecimento satisfatório. A rotina era composta das seguintes tarefas: visualizar na tela os campos apontados como não-reconhecidos e digitar as informações desses campos;

- Crítica – destinava-se à validação da consistência de alguns dos dados capturados e dos totais de questionários de um setor através das seguintes verificações: do número de questionários pertencentes ao mesmo setor; da seqüência de questionários por setor; e do total de pessoas por questionário;

- Transmissão – os arquivos de dados, resultantes da passagem dos questionários pelo processo de produção, eram transmitidos para o Centro de Processamento de Dados do Rio de Janeiro; e

- Rotinas de Segurança – tratava das rotinas de segurança da gravação dos dados nas mídias determinadas – CD-ROM e fitas DAT.

A organização, o funcionamento e os processos de trabalho foram testados no Centro de Captura do Rio de Janeiro, instalado em março de 2000, para a apuração do Censo Experimental. Entre os testes realizados, destacam-se:

- os aplicativos desenvolvidos;

- os parâmetros de crítica e verificação;

- as taxas de reconhecimento de caracteres numéricos, alfanuméricos, marcas e código de barras:

- os relatórios de acompanhamento da produção;

- as formas de atuação dos coordenadores e supervisores;

- as rotinas de manutenção dos equipamentos e, em especial, dos scanners;

- a melhor forma de retirada dos grampos dos questionários da amostra e a escolha da melhor lâmina para o corte dos mesmos;

- o treinamento dos coordenadores e supervisores dos outros CCDs nas rotinas e nos aplicativos;

- o treinamento das equipes de verificação e crítica;

- o treinamento dos operadores de *scanners*;

- as rotinas de segurança dos dados, incluindo *back-up* dos servidores e das imagens dos questionários;

- as rotinas de gravação dos CDs e transmissão dos dados para o *mainframe*;

- as soluções para o descarte das pastas após a digitalização;

- as rotinas de recepção, identificação, preparação e armazenamento das pastas.

3.5 Impressão, distribuição e retorno do material

Uma vez definidos os instrumentos a serem utilizados no Censo 2000, houve a fase de impressão. Depois do material impresso, começou a distribuição, uma etapa fundamental dentro da logística de uma operação censitária. De posse dos instrumentos, os Recenseadores foram a campo e realizaram a pesquisa, sendo o retorno do material preenchido parte da logística do Censo 2000. Cada etapa citada nesta introdução está detalhada a seguir.

3.5.1 Impressão

O material impresso, utilizado no Censo 2000, pode ser dividido em três grupos: instrumentos de treinamento, instrumentos de coleta e instrumentos administrativos. A impressão desse material foi realizada na gráfica do IBGE, com exceção dos Questionários Básico e da Amostra, conforme será visto a seguir.

Em linhas gerais, a impressão do material que atendeu às atividades de treinamento dos Recenseadores alcançou, nos meses de maio e junho de 2000, a tiragem planejada, havendo atraso apenas na produção do Manual do Recenseador. Já a impressão do material que atendeu à coleta teve a tiragem planejada alcançada em julho de 2000.

Tabela 3.5 - Número de exemplares impressos para o Censo Demográfico, segundo o grupo de instrumento - 2000

Grupo de instrumentos	Número de exemplares
Total	**89 991 250**
Instrumentos de treinamento	3 783 250
Instrumentos de coleta	79 670 000
Instrumentos administrativos	6 538 000

Fonte: IBGE, Centro de Documentação e Disseminação de Informações.

As maiores tiragens planejadas para os materiais de coleta do Censo foram orientadas tomando-se por base a estimativa de domicílios para o ano de 2000 – equivalente a 45 milhões. A tabela a seguir mostra o total de instrumentos impressos e o número de exemplares produzidos por grupo de material.

3.5.1.1 Instrumentos de treinamento

As inovações pedagógicas do treinamento do Censo 2000 tiveram reflexo no número de instrumentos aplicados, que foi significativamente superior ao de operações anteriores. O material foi ilustrado, diagramado, editorado e impresso integralmente no IBGE.

A tabela a seguir relaciona os instrumentos de treinamento utilizados no Censo 2000 e a quantidade produzida de cada um deles.

Tabela 3.6 - Número de exemplares impressos para o treinamento do Censo Demográfico, segundo o tipo de instrumento - 2000

Instrumentos de treinamento		Número de exemplares
	Total	**3 783 250**
CD-1.01	Questionário básico	300 000
CD-1.02	Questionário da amostra	600 000
CD-1.03	Folha de domicílio coletivo	300 000
CD-1.06	Caderneta do setor	40 000
CD-1.07	Folha de coleta	600 000
CD-1.09	Manual do recenseador	300 000
CD-1.10	Manual do supervisor	40 000
CD-1.11	Manual do agente censitário municipal	13 000
CD-1.12	Manual de coordenação técnica	3 500
CD-1.15	Agenda do supervisor	40 000
TR-01	Manual do instrutor	40 000
TR-02	Roteiro de estudo para o recenseador	300 000
TR-03	Teste inicial	300 000
TR-04	Caderno de exercícios	300 000
TR-05	Teste final G1	440
TR-06	Teste final A	40 100
TR-07	Teste final B	260 000
TR-08	Gabarito do teste final G1	150
TR-09	Gabarito do teste final A	3 500
TR-10	Gabarito do teste final B	22 000
TR-11	Avaliação do treinamento - treinando	3 500
TR-12	Avaliação do treinamento - instrutor	2 000
TR-13	Avaliação do treinamento - observador	700
TR-14	Álbum seriado	13 000
TR-15 A	Instrumentos ampliados - folha de coleta	32 000
TR-15 B	Instrumentos ampliados - folha de coleta	22 000
TR 16	Instrumentos ampliados - folha de domicílio coletivo	22 000
TR-17	Roteiro das aulas para treinamentos de 12 horas	8 000
TR-18	Manual do instrutor	12 000
TR-19	Caderno de exercícios - base operacional geográfica	42 000
TR-20	Álbum seriado II	9 000
TR-21	Avaliação do treinamento - treinando	1 800
TR-22	Avaliação do treinamento – instrutor	800
TR-23	Instruções para revisão do trabalho da base operacional geográfica	40 000
TR-24	Manual do usuário - sistema indicadores gerenciais da coleta de dados	40 000
TR-25	Vídeo Censo 2000 treinamento	12 700
TR-26	Vídeo Censo 2000 institucional	660
TR-29	Mapa - MME	9 000
TR-30	Mapa - MLE	9 000
TR-31	Avaliação do treinamento - observador	400

Fonte: IBGE, Centro de Documentação e Disseminação de Informações.

3.5.1.2 Instrumentos de coleta

Os instrumentos de coleta tiveram seu conteúdo definido pelas áreas técnicas e depois receberam a formatação no CDDI. Para a impressão dos questionários, foi realizada licitação, tendo em vista que se tratava de impressão com dados variáveis, em grande volume e considerando que cada folha deveria receber numeração – número do questionário e da folha.

Inicialmente, pensava-se em produzir o Questionário Básico na gráfica do IBGE e o da Amostra em gráfica externa, visto que seu formato exigia impressão em máquina rotativa. Foram contatados os representantes dos fabricantes

de equipamentos de impressão por jato de tinta e empresas que utilizavam essa tecnologia. A gráfica digital do IBGE já possuía equipamentos e domínio da tecnologia de códigos de barras, mas para pequenas e médias tiragens, que não atendiam as exigências do Censo. Além disso, um dos pontos mais delicados era a manutenção desse tipo de equipamento, em virtude do número reduzido de empresas que o utilizavam. Este motivo, somado a outros fatores, como o alto custo dos equipamentos e a pouca visibilidade de uso futuro, desencorajou a sua compra. Desta forma, todos os instrumentos com impressão de dados variáveis foram produzidos externamente.

Os instrumentos de coleta foram elaborados de modo que atendessem às necessidades do Recenseador, priorizando características como a facilidade de leitura e registro, auxiliando o trabalho de campo, e adaptado à digitalização e reconhecimento ótico de caracteres. A gramatura do papel foi de 90g de alta alvura e opacidade, e a cor da tinta especificada foi azul, escala Pantone 313 U.

Durante a produção, foi implantado um sistema de controle de qualidade, no qual, sistematicamente, eram retiradas e enviadas para teste, amostras dos questionários produzidos. Não foram detectadas falhas significativas pelo controle de qualidade, nem mais tarde, na fase de captura dos dados dos questionários.

Com relação às tiragens – apresentadas na próxima tabela – vale destacar algumas observações referentes aos questionários e às Folha de Coleta. O Questionário Básico teve tiragem diferenciada para grupos de municípios, considerando a estimativa do total de domicílios para o ano 2000. A definição das quantidades a serem impressas foi feita da seguinte forma:

- para municípios com menos de 15 mil habitantes: 96% da estimativa;

- para municípios com 15 mil até 100 mil habitantes: 108% da estimativa; e

- para municípios com mais de 100 mil habitantes:

  108% da estimativa, para os municípios da Região Norte - exceto Manaus e Belém;e
  90% da estimativa, para os demais municípios brasileiros.

Seguindo o cálculo diferenciado da tiragem, o Questionário da Amostra, para as 5,5 milhões de cópias previstas, alcançou as seguintes parcelas da estimativa do número de domicílios:

- 24%, para os municípios com menos de 15 mil habitantes;

- 12%, para os municípios com 15 mil até 100 mil habitantes; e

- 10%, para os municípios com mais de 100 mil habitantes.

Já a Folha de Coleta teve uma tiragem correspondente a 200% da calculada para o Questionário da Amostra.

É importante registrar que, durante a etapa de coleta, para atender às demandas adicionais das unidades regionais do IBGE, foi necessário produzir uma tiragem complementar do Questionário Básico – mais 28% – do Questionário da Amostra – mais 36% – e da Folha de Coleta – mais 16%.

Na tabela a seguir, estão relacionados os instrumentos de coleta produzidos para o Censo 2000 e suas respectivas tiragens.

Tabela 3.7 - Número de exemplares impressos para a coleta do Censo Demográfico, segundo o tipo de instrumento - 2000

Instrumentos de coleta		Número de exemplares
	Total	**79 670 000**
CD-1.01	Questionário básico	55 000 000
CD-1.02	Questionário da amostra	8 000 000
CD-1.03	Folha de domicílio coletivo	600 000
CD-1.06	Caderneta do setor	250 000
CD-1.07	Folha de coleta	15 000 000
CD-1.13	Guia de remessa de questionários	50 000
CD-1.14	Guia de remessa das folhas de coleta	50 000
CD-1.16	Registro dos setores iniciados	120 000
CD-1.17	Resumo da produção semanal	600 000

Fonte: IBGE, Centro de Documentação e Disseminação de Informações.

3.5.1.3 Instrumentos administrativos

Os números de exemplares impressos dos instrumentos específicos para a administração da operação censitária, no que se refere aos recursos humanos, materiais, rotinas e procedimentos de contratação, pagamento, entre outros – instrumentos administrativos – constam da tabela seguinte.

Tabela 3.8 - Número de exemplares impressos para as atividades administrativas do Censo Demográfico, segundo o tipo de instrumento - 2000

(continua)

Instrumentos administrativos		Número de exemplares
	Total	**6 538 000**
AD-01	Manual do sistema de administração de pessoal - orientações gerais	44 500
AD-02	Manual do sistema de administração de pessoal - cadastro	44 500
AD-03	Manual do sistema de administração de pessoal - pagamento	2 000
AD-04	Ficha de cadastro	775 000
AD-05	Ficha de exclusão do cadastro	13 000
AD-06	Requerimento de dependência - imposto de renda	130 000
AD-07	Requerimento de salário-família	130 000
AD-08	Declaração de dependência econômica	91 000
AD-09	Declaração Lei nº 8 745	323 000
AD-10	Declaração de não-acumulação de cargos públicos	323 000
AD-11	Programa de assistência ao censitário	323 000
AD-12	Contrato e termo de responsabilidade do ACM	20 000
AD-13	Contrato e termo de responsabilidade do supervisor	70 000
AD-14	Contrato e termo de responsabilidade do recenseador	560 000
AD-15	Crachá	360 000
AD-16	Ficha de contratação	775 000

Tabela 3.8 - Número de exemplares impressos para as atividades administrativas do Censo Demográfico, segundo o tipo de instrumento - 2000

(conclusão)

Instrumentos administrativos		Número de exemplares
AD-17	Ficha de convocação	180 000
AD-18	Termo de desistência	90 000
AD-19	Requerimento de desligamento	90 000
AD-20	Comunicação de desligamento	140 000
AD-21	Ficha de desligamento	97 000
AD-22	Pensão alimentícia	29 000
AD-23	Solicitação do auxílio-transporte recenseador	70 000
AD-24	Solicitação do auxílio-transporte mensalista	43 000
AD-25	Folha de freqüência	390 000
AD-26	Folha de afastamento	21 000
AD-27	Termo aditivo (ACM/ACS)	6 000
AD-28	Prorrogações	600 000
AD-29	Pagamento do adiantamento da produção	48 000
AD-30	Solicitação de pagamento por produção	250 000
AD-31	Solicitação de pagamento ajuda de custo transporte	250 000
AD-32	Contrato e termo de responsabilidade ACA	250 000

Fonte: IBGE, Centro de Documentação e Disseminação de Informações.

3.5.2 Separação, embalagem e distribuição

A distribuição do material foi planejada para ocorrer a partir da gráfica do IBGE, em Parada de Lucas, no Rio de Janeiro, para as agências de todo o país, que então enviariam o material aos postos de coleta. Um centro de distribuição, funcionando em um galpão, foi instalado para a execução das etapas de separação e embalagem do material.

Na época, a empresa que prestava serviços ao IBGE foi a responsável pelo transporte, tendo recebido as planilhas com a relação de locais e os quantitativos para planejar previamente a distribuição. Nesta fase, vale destacar que:

- a área total do galpão em Parada de Lucas era de 2 400 m^2, dos quais foram destinados 1 800 para a distribuição e circulação;

- a empresa transportadora deveria estudar e planejar os roteiros de entregas; e

- a separação e embalagem de material exigia alta intervenção manual, visto que, para a maior parte dos instrumentos remetidos, as quantidades representavam frações do conteúdo de uma caixa.

Com o início da remessa do material para a fase de treinamento dos Recenseadores, os problemas começaram a aparecer. O espaço era insuficiente para o funcionamento de duas linhas de separação, encaixotamento e estocagem do material, assim como a área de manobra e carga não comportava o volume de mais de dois veículos simultaneamente. Além do que, duas linhas de encaixotamento eram insuficientes para dar um ritmo acelerado à expedição do

material. Enquanto isso, o IBGE era informado de que a empresa transportadora estava entregando as cargas, obedecendo à ordem alfabética dos municípios das agências, dentro de cada Unidade da Federação. As medidas adotadas frente a esses problemas foram:

- aumentar para três turnos o trabalho de expedição e estabelecer quatro linhas de separação e encaixotamento;

- assumir a confecção dos roteiros de entrega e os despachos;

- dar preferência a caminhões de menor porte, por serem mais ágeis para manobra, assim como mais rápidos em atingir os destinos; e

- exigir cumprimento de prazos da transportadora – parte da carga foi expedida por via aérea.

A esses problemas, foi acrescido o atraso no fornecimento das pastas do Recenseador.

Na distribuição inicial do material de coleta do Censo, ocorrida ao longo dos meses de junho e julho de 2000, utilizaram-se transporte aéreo e terrestre, com chegadas previstas para até um mês aproximadamente. O transporte terrestre foi acompanhado através do contato telefônico dos motoristas, com o IBGE, por meio do 0800-218181, dando notícias das chegadas aos diversos municípios.

Houve uma preocupação maior com os questionários, que passaram pela impressão de dados variáveis. Esse material recebeu uma identificação diferenciada no encaixotamento, de modo que fosse possível rastreá-lo. Desta forma, se houvesse defeito em um determinado lote, era possível saber que unidades receberam esses questionários, localizando o material que apresentava problemas.

Foi necessária também uma nova remessa de alguns instrumentos durante a etapa de coleta, para atender às demandas adicionais das Unidades Regionais do IBGE. Foram enviados Questionários Básico e da Amostra e Folhas de Coleta, em proporções correspondentes aos acréscimos de impressão já citados.

A distribuição das pastas do Recenseador, bem como de lapiseiras, borrachas, canetas etc. aconteceu nos meses de julho e agosto de 2000. Foram distribuídas, aproximadamente: 213 mil pastas, 247 mil lapiseiras, 295 mil caixas de grafites, 247 mil borrachas, 199 mil canetas e 198 mil pranchetas.

3.5.3 Retorno do material coletado e armazenamento

Ainda dentro da logística da operação censitária, estão duas etapas pós-coleta: o retorno do material e o armazenamento. Em linhas gerais, o material preenchido – questionários, Cadernetas de Setor e Folhas de Coleta – parte permaneceu nas Unidades Regionais e parte seguiu para os Centros de Captura de Dados, onde seria iniciada a fase da apuração dos dados. Nos CCDs, o material era armazenado até o seu processamento, e após a captura dos dados, o material ainda permaneceu armazenado por um determinado período.

Os itens a seguir descrevem, detalhadamente, o caminho percorrido pelo material preenchido até os CCDs.

3.5.3.1 Remessas

Uma vez encerrada a coleta de dados no município, coube ao Agente Censitário Municipal a organização e o acondicionamento dos questionários, Cadernetas de Setor e Folhas de Coleta, segundo as orientações que constavam no Manual do Agente Censitário Municipal, bem como o seu encaminhamento ao Coordenador de Subárea, normalmente sediado em uma agência do IBGE. Posteriormente, de acordo com as orientações da Divisão Estadual de Pesquisas, parte do material era remetido para a capital a fim de que fosse então enviado para o Centro de Captura de Dados, responsável pelo processamento desses instrumentos. As remessas de material das unidades regionais para os CCDs eram estabelecidas entre os titulares das unidades envolvidas e levavam em conta a capacidade de armazenamento e a programação de processamento de cada CCD.

Para preservar a integridade física dos documentos que seriam digitalizados – questionários e Folhas de Coleta – durante os deslocamentos, foram utilizadas duas espécies de caixa de papelão: uma para os questionários e Folhas de Coleta e outra, maior, chamada caixa máster, para acondicionamento da primeira. Na caixa pequena, cabiam cerca de 300 Questionários Básicos e 35 da Amostra, ou seja, cada caixa comportava, em geral, o material de um setor, separado por modelo de questionário. No caso da Folha de Coleta, a caixa comportava o material de vários setores, variando segundo a quantidade de páginas de cada setor.

Os questionários de cada setor eram empastados separadamente, segundo seu tipo, de forma que cada setor resultava em duas caixas – uma com os Questionários Básicos e outra com os Questionários da Amostra. A chave de identificação do setor – UF/Município/Distrito/Subdistrito/Setor – o modelo do questionário e a numeração da caixa – 1/2 ou 2/2, por exemplo – eram registrados nas quadrículas pré-impressas na lombada das caixas pequenas. Cada caixa máster comportava seis caixas pequenas, e era também numerada para fins de controle da remessa.

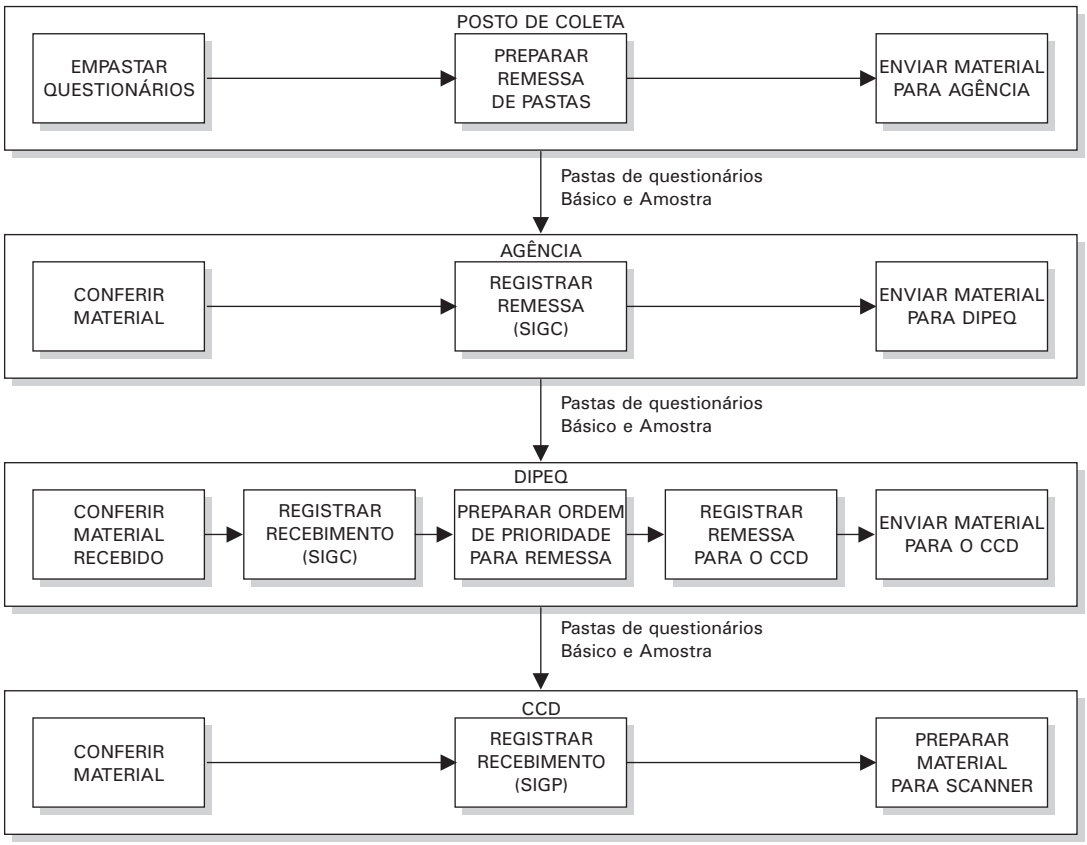
Cada remessa feita pelo posto de coleta era precedida do preenchimento de uma guia – Guia de Remessa de Questionários ou a Guia de Remessa de Folhas de Coleta – nas quais eram relacionados os setores encaminhados. Os postos informatizados dispunham de guia eletrônica, e os demais utilizavam formulário impresso, para posterior emissão de guia eletrônica na agência ou posto informatizado, sede do Coordenador de Subárea.

Embora fosse possível a remessa parcial de setores entre os postos de coleta e as agências, e destas para as sedes estaduais, de modo geral isso não ocorreu, já que os ACMs preferiram fazer remessas completas, ou seja, de todos os questionários do município ou de sua área de atuação.

A seguir, é apresentado o fluxo do material coletado, desde o posto de coleta até o preparo para o escaneamento no CCD.

Figura 3.5 – Fluxo de remessa de questionários coletados

Do Posto de Coleta ao Centro de Captura de Dados



3.5.3.2 Armazenamento

Inicialmente, não foi planejado o armazenamento de grandes volumes de material nas unidades regionais. O programado era que as agências, a partir do início do mês de outubro de 2000, fizessem remessas parciais de material para as capitais, onde seriam formados lotes de dimensões adequadas para o envio aos respectivos CCDs, segundo uma programação previamente combinada. Contudo, o atraso na normalização da operação dos Centros de Captura de Dados, inicialmente prevista para o mês de outubro de 2000 e efetivamente ocorrida em fevereiro de 2001, implicou a retenção do material nos estados, gerando a necessidade de disponibilização de espaços de armazenamento não-previstos.

O armazenamento de todo o material processado nos CCDs também ocorreu sem que estivesse programado. A intenção inicial era de que lotes de questionários fossem destruídos, tão logo os arquivos de dados correspondentes fossem recebidos e validados pela unidade responsável, situada no Rio de Janeiro. Desta forma, os espaços reservados em cada CCD para o armazenamento de questionários processados, comportavam apenas 20% a 30% do total de material esperado. No entanto, a Diretoria de Pesquisas, dadas algumas incertezas inerentes ao uso da nova tecnologia de captura de dados, não autorizou o descarte dos questionários, até que fossem feitos diversos testes e avaliações nos arquivos de dados e nas imagens digitalizadas. Assim, os questionários somente puderam ser destruídos após o encerramento da etapa de captura de

dados, o que implicou a guarda de 100% do material processado em cada CCD. Essa situação gerou a necessidade de ampliação de espaços com novas locações, cessões ou a ocupação de outras áreas.

3.5.4 Destruição dos questionários

A Diretoria de Pesquisas autorizou a destruição dos questionários somente depois de ter sido dada como concluída a etapa de apuração dos dados em todos os Centros de Captura de Dados, quando todos os arquivos de dados haviam sido recebidos e validados no Centro de Processamento de Dados do IBGE no Rio de Janeiro e todas as imagens geradas se encontravam devidamente gravadas nas mídias determinadas – CD-ROM e fitas DAT.

Para a alienação dos questionários foi necessário montar um esquema operacional cercado de cuidados que garantisse a manutenção do sigilo sobre o conteúdo de cada um deles. Dado o grande volume de material e a necessidade de envolvimento de outras empresas e de pessoas externas ao IBGE, foram estabelecidas as condições básicas que orientaram a escolha das empresas responsáveis pela tarefa de alienação em cada município sede de CCD, quais sejam:

- que todo o material retirado do local de armazenamento em um dia fosse destruído no mesmo dia, obedecendo ao horário comercial;

- que todas as fases do processo – retirada, pesagem, transporte, carga/descarga e destruição – seriam acompanhadas por comissão de servidores do IBGE especialmente designada; e

- que, preferenciamente, as empresas concorrentes à prestação do serviço, fossem estabelecidas nos municípios sede de cada um dos CCDs, com vistas à redução de custos.

Tomando em conta essas regras, foram levantados os cadastros e visitados os estabelecimentos que negociavam com aparas e/ou com fabricação de papel em cada um dos municípios sede dos CCDs. Essas visitas tinham o objetivo de reunir subsídios para a elaboração do edital de licitação, de forma a garantir que, pelo menos, três empresas tivessem condições para participar da concorrência em cada município. Na ocasião, foram também obtidas informações a respeito dos preços praticados pelo mercado para material semelhante ao que seria alienado.

No curso dos estudos para a alienação do material, o Prefeito de Campina Grande/PB solicitou ao IBGE a doação dos questionários armazenados no CCD daquela cidade, a fim de aplicar os recursos provenientes de sua venda no Fundo Municipal de Assistência à Criança e ao Adolescente. Após exame do referido pedido pelos órgãos competentes do IBGE, e considerando que aquela Prefeitura havia cedido gratuitamente o imóvel em que funcionou o CCD, a Direção do IBGE decidiu atender ao pedido de doação, mediante assinatura de termo de doação. Embora a Prefeitura tenha ficado com a responsabilidade da escolha da empresa encarregada da destruição do material, todos os ritos da operação seguiram o modelo fixado pelo IBGE, inclusive o acompanhamento da comissão de servidores.

Assim, foram realizadas as licitações, na modalidade de carta-convite, para alienação do material – questionários e caixas de papelão – dos CCDs de Campinas, Curitiba, Goiânia e Rio de Janeiro. Ao todo, incluído o material doado à Prefeitura de Campina Grande, foram alienadas mais de 1 000 toneladas de material.

Referências

AVALIAÇÃO dos censos 96: relatório final. Rio de Janeiro: IBGE, Diretoria de Planejamento e Coordenação, 1998. 28 p.

BRASIL. Decreto-lei nº 969, de 21 de dezembro de 1938. Dispõe sobre os recenseamentos gerais do Brasil. *Coleção de leis [da] República Federativa do Brasil*, Brasília, DF, v. 4, p. 315, 1938. Col. 1.

_____. Lei nº 7.853, de 24 de outubro de 1989. Dispõe sobre o apoio às pessoas portadoras de deficiência, sua integração social, sobre a Coordenadoria Nacional para Integração da Pessoa Portadora de Deficiência - CORDE, institui a tutela jurisdicional de interesses coletivos ou difusos dessas pessoas, disciplina a atuação do Ministério Público, define crimes, e dá outras providências. Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF, 25 out. 1989. p. 1920. Col. 1.

_____. Lei nº 8.666, de 21 de junho de 1993. Regulamenta o artigo 37, inciso XXI, da Constituição Federal, institui normas para licitações e contratos da Administração Pública e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 22 jun. 1993. p. 8269. Col. 1. Retificada no Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF, 2 jul. 2003. p. 1. Col. 3.

_____. Lei nº 8.745, de 9 de dezembro de 1993. Dispõe sobre a contratação por tempo determinado para atender a necessidade temporária de excepcional interesse público, nos termos do inciso IX do artigo 37 da Constituição Federal, e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 10 dez. 1993. p. 18937. Col. 2.

_____. Decreto nº 2.424, de 17 de dezembro de 1997. Autoriza excepcionalmente a contratação temporária de recursos humanos para os órgãos e entidades do Poder Executivo Federal que especifica. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 18 dez. 1997. p. 30225. Col. 2.

_____. Ministério da Administração Federal e Reforma do Estado. Despacho de 10 de março de 1998. Autoriza a Fundação Instituto Brasileiro de Geografia e Estatística – IBGE a contratar 700 (setecentas) pessoas para atender necessidade temporária de excepcional interesse público. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, n. 47, 11 mar. 1998. Seção 1, p. 128. Col. 1.

_____. Medida provisória nº 1.748-40, de 8 de abril de 1999. Altera os artigos 2, 3, 4, 5, 6, 7 e 9 da Lei nº 8.745, de 9 de dezembro de 1993, que dispõe sobre a contratação por tempo determinado para atender a necessidade temporária de excepcional interesse público, e dá outra providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 8 abr. 1999. p. 11. Col. 1.

_____. Secretaria de Estado da Administração e do Patrimônio. Ofício nº 283 SEAP/MOG, de 11 de maio de 1999. Autoriza a Fundação Instituto Brasileiro de Geografia e Estatística – IBGE a contratar até 250 (duzentos e cinqüenta) profissionais para as atividades relacionadas à coleta de dados do Censo Experimental. Brasília, DF, 1999.

AVALIAÇÃO dos censos 96: relatório final. Rio de Janeiro: IBGE, Diretoria de Planejamento e Coordenação, 1998. 28 p.

_____. Ministério do Planejamento, Orçamento e Gestão. Portaria nº 127, de 21 de setembro de 1999. Autoriza o Instituto Brasileiro de Geografia e Estatística – IBGE a contratar até 225 030 pessoas para a realização do Censo Demográfico 2000, para atender a necessidade temporária de excepcional interesse público. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, n. 182, 22 set. 1999. Seção 1, p. 21. Col. 1.

_____. Lei nº 9.849, de 26 de outubro de 1999. Altera os artigos 2º, 3º, 4º, 5º , 6º, 7º e 9º da Lei nº 8.745, de 9 de dezembro de 1993, que dispõe sobre a contratação por tempo determinado para atender a necessidade temporária de excepcional interesse público, e dá outra providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 27 out. 1999. Seção Extra, p. 4. Col. 2.

_____. Decreto nº 3.272, de 3 de dezembro de 1999. Aprova o estatuto e o quadro demonstrativo dos cargos em comissão e funções gratificadas da Fundação Instituto Brasileiro de Geografia e Estatística – IBGE, e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 6 dez. 1999. p. 4. Col. 1.

_____. Decreto nº 3.298, de 20 de dezembro de 1999. Regulamenta a Lei nº 7.853, de 24 de outubro de 1989, dispõe sobre a política nacional para a integração da pessoa portadora de deficiência, consolida as normas de proteção, e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 21 dez. 1999. p. 10. Col. 1.

_____. Ministério do Planejamento, Orçamento e Gestão. Portaria nº 125, de 15 de junho de 2000. Concede suprimentos de fundos especiais, no valor máximo de R$ 10.000,00 (dez mil reais), por suprimento, com prazo de aplicação de até 60 (sessenta) dias cada, para atendimento de despesas de pronto pagamento referentes aos trabalhos de Contagem Populacional e Censo Agropecuário, cada pagamento limitado ao teto legal (art. 24, inciso II, da Lei nº 8666/93) e sujeito à posterior prestação de contas (par. 2º do art. 45, do Decreto nº 93.872/86). *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, n. 116, 16 jun. 2000. Seção 1, p. 61. Col. 1.

CENSO demográfico 2000: manual do agente censitário municipal – CD 1.11. Rio de Janeiro: IBGE, 2000. 33 p.

_____: manual do recenseador – CD 1.09. Rio de Janeiro: IBGE, 2000. 151 p.

FELLEGI, I. P.; HOLT, D. A systematic approach to automatic edit and imputation. *Journal of the American Statistical Association*, Washington, D.C., v. 71, n. 353, p. 17-35, 1976.

IBGE. Resolução do Conselho Diretor - CD nº 0006/97, de 12 de junho de 1997. Institui Comissão de Planejamento e Organização Geral do Censo 2000. Rio de Janeiro, 1997. Disponível em: <http://w3.aplicacao.de.ibge.gov.br/Siad/SIAD.nsf>. Acesso em: out. 2003.

_____. Resolução do Conselho Diretor - CD nº 0015/99, de 7 de dezembro de 1999. Aprova estrutura organizacional do IBGE. Rio de Janeiro, 1999. Disponível em: <http://w3.aplicacao.de.ibge.gov.br/Siad/SIAD.nsf>. Acesso em: out. 2003.

IBGE. Resolução do Conselho Diretor - CD nº 0006/97, de 12 de junho de 1997. Institui Comissão de Planejamento e Organização Geral do Censo 2000. Rio de Janeiro, 1997. Disponível em: <http://w3.aplicacao.de.ibge.gov.br/Siad/SIAD.nsf>. Acesso em: out. 2003.

_____. Resolução do Conselho Diretor - CD nº 003/00, de 1 de março de 2000. Concede gratificação por serviços extraordinários para as atividades do Censo Demográfico de 2000. Rio de Janeiro, 2000. Disponível em: <http://w3.aplicacao.de.ibge.gov.br/Siad/SIAD.nsf>. Acesso em: out. 2003.

_____. Resolução do Presidente - PR nº 0001/00, de 1 de março de 2000. Constitui os Centros de Captura de Dados - CCDs, com a finalidade de executar as atividades relativas à entrada de dados dos questionários do Censo 2000, por meio de reconhecimento ótico, nas cidades de Campina Grande - PB, Campinas - SP, Curitiba - PR, Goiânia - GO e Rio de Janeiro - RJ. Rio de Janeiro, 2000. Disponível em: <http://w3.aplicacao.de.ibge.gov.br/Siad/SIAD.nsf>. Acesso em: out. 2003.

MANUAL de apoio operacional: censo 2000. Rio de janeiro: IBGE, Coordenação de Orçamento e Finanças, 2000. 79 p.

NORMAS de apresentação tabular. 3. ed. Rio de Janeiro: IBGE, 1993. 62 p.

PROJETO centro de captura de dados para o Censo 2000. Rio de Janeiro: IBGE, Diretoria de Informática, 1998. 11 p.

RELATÓRIO de avaliação da etapa de coleta do censo 2000. Rio de Janeiro: IBGE, Diretoria Executiva, 2001. 22 p.

4 Mobilização

Para que uma operação do porte do Censo Demográfico 2000 fosse realizada com êxito, além de recursos financeiros, materiais e humanos, era necessária também a mobilização da população. Era fundamental que os diversos segmentos da sociedade estivessem cientes da existência da pesquisa, da sua importância e, sobretudo, do seu papel como informante. Para alcançar esse objetivo, foram desenvolvidas ações gerais e outras voltadas para grupos específicos, detalhadas a seguir.

4.1 Concurso da logomarca

O IBGE organizou um concurso para a escolha da logomarca do Censo Demográfico 2000, como ação preparatória de mobilização da população, uma vez que o símbolo identificaria todo o material utilizado no Censo, desde os questionários até as peças publicitárias.

O objetivo era identificar, imediatamente, através de uma imagem, o Censo 2000. A cada anúncio, cartaz, folheto, *outdoor*, ou qualquer outro material com a logomarca, a realização do Censo seria lembrada por quem a visse.

O edital do concurso, lançado em julho de 1998, foi dirigido a alunos das faculdades de comunicação visual, desenho industrial, propaganda e publicidade, artes e correlatas, de todo o País. Formou-se uma comissão julgadora, integrada por oito membros, sendo cinco externos ao IBGE, para selecionar as 30 melhores logomarcas entre 916 concorrentes.

Os membros da comissão julgadora do concurso da logomarca do Censo 2000 foram:

Membros externos

- Angela Carvalho – *Designer*, professora da Coordenação dos Programas de Pós-graduação em Administração – COPPEAD – da Universidade Federal do Rio de Janeiro – UFRJ;

- João de Souza Leite – Vice-Diretor da Escola Superior de Desenho Industrial – ESDI – da Universidade Estadual do Rio de Janeiro – UERJ;

- Meire Marin Hernandes – Coordenadora de Redação da Revista Design Gráfico, São Paulo;

- Pojucan – Ilustrador do *Jornal do Brasil* e integrante da produção do programa Casseta e Planeta Urgente, TV Globo;

- Sebastião Amoedo – Coordenador do Curso de Publicidade e Propaganda da Universidade Federal do Rio de Janeiro – UFRJ

Membros internos

- David Wu Tai – Superintendente do Centro de Documentação e Disseminação de Informações – CDDI – IBGE;

- Maria Vilma Salles Garcia – Coordenadora da Coordenação de Acompanhamento e Controle Operacional dos Censos – IBGE; e

- Aldo Victorio Filho – Chefe da Divisão de Criação do Centro de Documentação e Disseminação de Informações – IBGE

Dos 30 trabalhos selecionados, três foram premiados:

- 1º lugar: Henrique Del Bosco Loyola Borges – Pontifícia Universidade Católica do Paraná – PUC-PR – Desenho Industrial: Prêmio: R$ 5.000,00;

- 2º lugar: Bruno Basso – Universidade Santa Cecília, Santos/SP – Publicidade e Propaganda: Prêmio: R$ 2.500,00; e

- 3º lugar: Daniela Wajman – Faculdade de Belas Artes de São Paulo – Desenho Industrial: Prêmio: R$ 1.000,00.

Os 30 melhores trabalhos, com a identificação da autoria, foram publicados na página do IBGE, na Internet. Os autores tiveram seus nomes publicados no Diário Oficial da União, de 20 de outubro de 1998. A logomarca vencedora do concurso, apresentada a seguir, foi adotada pelo IBGE, passando a identificar o Censo Demográfico 2000 em todas as peças promocionais, de papelaria, e documentos relacionados à pesquisa.

Figura 4.1 - Logomarca do Censo 2000

Essa imagem foi considerada, pelos jurados, de fácil entendimento, por remeter a uma das formas mais simplificadas de contagem e pela associação desse conceito à operação do Censo 2000; atendia também a critérios técnicos, como imagem impactante e de fácil reprodução em todos os materiais.

4.2 Projeto Vamos Contar!

Vamos Contar! foi um projeto didático-pedagógico, planejado e elaborado durante os anos de 1999 e 2000, que teve como objetivo envolver alunos e professores das escolas do Ensino Fundamental e do Ensino Médio, das redes pública e privada do País, no esforço de mobilização para o Censo 2000. Através do envio de material didático para as escolas, preparado especialmente para essa finalidade, o IBGE pretendia criar um ambiente favorável ao recebimento dos Recenseadores pela população, e divulgar a importância do censo e dos seus resultados.

4.2.1 Histórico

A primeira proposta de um projeto envolvendo as escolas data do Censo de 1970. Na ocasião, foi distribuída aos alunos do Curso Primário – atual 1ª a 4ª série do Ensino Fundamental – uma revista intitulada *"Julinho, o recenseador"*, que contava, em forma de estória em quadrinhos, a importância da realização do censo. A revista fazia parte do plano de propaganda do Recenseamento Geral da República, com o objetivo de criar, entre as crianças, um movimento de interesse pela operação censitária.

Para a campanha do Censo Demográfico 1991, foi criado um *folder*, com o nome de Projeto Escola, contendo informações sobre a operação da pesquisa. Durante a realização do censo, a comunicação, embutida no material censitário, foi entregue às escolas públicas, alcançando grande sucesso.

A partir dessas experiências, cresceu a necessidade de um projeto com maior envolvimento da comunidade escolar. Essa foi a proposta do Projeto Vamos Contar!

A idéia foi elaborar um projeto em "duas mãos". Por um lado, haveria o envolvimento das escolas na campanha do Censo 2000. Por outro, o IBGE levaria às escolas material didático que serviria de apoio pedagógico aos professores e de instrumental para as atividades em sala de aula.

Em março de 1999, foi elaborado um estudo para definir o objetivo do projeto, o material envolvido e o público que se pretendia alcançar. O material produzido pelo Scholastic/U.S. Census Bureau, dos Estados Unidos, uma iniciativa semelhante a que se pretendia com o Vamos Contar!, constituiu a base do projeto.

A proposta foi envolver as escolas do Ensino Fundamental e do Ensino Médio das redes pública e privada do país, alcançando, aproximadamente, 200 000 escolas em todo o Brasil.

Objetivando o envolvimento dos diferentes atores relacionados com a escola brasileira, o material elaborado teve como destino: diretores; professores; membros dos Conselhos Escola/Comunidade e de Associações de Pais e Mestres; e alunos, chegando, por estes últimos, até às comunidades e aos seus pais/responsáveis.

4.2.2 Objetivo

Com o estudo pronto, ficou definido que o projeto teria como objetivo convidar a comunidade escolar a participar da campanha de divulgação da realização do Censo 2000. Pretendeu-se, assim, criar uma força-tarefa nas escolas, que auxiliasse na conscientização das famílias e da comunidade sobre a importância do Censo 2000 para a população brasileira, propiciando um ambiente favorável ao recebimento dos Recenseadores durante a etapa de coleta de dados.

A idéia incluía aproveitar a oportunidade da realização do censo, para levar às escolas material e mapas, que serviriam de instrumental para o uso cotidiano dos professores em sala de aula. O material didático seria distribuído a cerca de 200 000 escolas das redes pública e privada do País, atingindo 1 milhão de salas de aula.

O projeto continha propostas e orientações de atividades que objetivavam subsidiar os professores no trabalho com mapas e com informações estatísticas, geográficas e cartográficas, para que os alunos pudessem melhor compreender a realidade brasileira, bem como os diversos usos que um país pode fazer dos resultados de um censo.

A efetiva participação de todos na escola, através das atividades sugeridas e desenvolvidas, permitiria o envolvimento dos pais/responsáveis, a partir do material e das informações que os alunos levariam para casa.

O IBGE avaliou que o momento do censo era favorável para trabalhar o tema cidadania com os alunos, auxiliando-os a perceber que colaborar com o Censo 2000 significava participar do processo democrático do País. Em uma linha de pensamento mais ampla, constituía uma oportunidade de cultivar o interesse pela informação estatística e geográfica entre os alunos, bem como de auxiliar na formação de futuros usuários das informações do IBGE, ainda quando crianças e jovens.

4.2.3 Material

Foram elaborados para o projeto os seguintes materiais e quantidades:

- Carta ao Diretor (200 000 exemplares);
- Carta ao Professor (1 550 000 exemplares);
- Carta aos Membros do Conselho Escola-Comunidade (200 000 exemplares);
- Carta aos Pais/Responsáveis, anexada nos guias dos professores;
- folheto com informações básicas (2 milhões e 300 mil de exemplares);
- folheto exclusivo do Programa Amigos da Escola (24 000 exemplares, produzidos pela TV Globo);
- cartaz (400 000 exemplares);
- Guia do Professor – Ensino Fundamental – 1ª a 4ª série (820 000 exemplares);
- Guia do Professor – Ensino Fundamental – 5ª a 8ª série (500 000 exemplares);

- Guia do Professor – Ensino Médio – 1ª a 3ª série (200 000 exemplares);

- Mapa Demográfico – Ensino Fundamental – 1ª a 4ª série (600 000 exemplares);

- Mapa Demográfico – Ensino Fundamental – 5ª a 8ª série e Ensino Médio (400 000 exemplares);

- Mapa Político – Ensino Fundamental e Ensino Médio, impresso no verso dos mapas acima citados (1 milhão de exemplares);

- Carta-resposta do Professor – Ficha de Avaliação do Projeto (1 550 000 exemplares);

- folheto explicativo de apresentação do projeto – mala direta (200 000 exemplares);

- vídeo – apresentação e divulgação do Projeto (5 600 cópias reproduzidas); e

- vídeos – 3 Módulos (1ª a 4ª série e 5ª a 8ª série do Ensino Fundamental e Ensino Médio) – explicação da proposta e conteúdo sugerido e exemplos de atividades – (5 600 cópias reproduzidas).

4.2.4 Público-alvo

O projeto buscava atingir 200 mil escolas do Ensino Fundamental e do Ensino Médio, das redes pública e privada do país, perfazendo 1 milhão de salas-de-aula. Estariam envolvidos, aproximadamente, 1 850 000 professores, 4 milhões de membros dos Conselhos Escola-Comunidade, compreendendo 43 milhões de alunos.

4.2.5 Equipes envolvidas

Para a implementação das atividades do projeto, foi necessário contar com o trabalho de consultores e de pessoal contratado, além da participação de técnicos do IBGE.

A coordenação do projeto ficou a cago de um técnico do IBGE, com dedicação integral desde abril de 1999. A partir de 2000, houve também a participação de dois técnicos – de nível superior e nível médio – além de outros ligados às áreas de criação e artes gráficas da Instituição.

Houve a contratação de pessoal especializado para preparar o conteúdo pedagógico do programa educacional, de acordo com a área específica. Assim, atuaram no projeto:

- um professor(a) responsável pelo conteúdo pedagógico do Ensino Fundamental da 1ª a 4ª série;

- um professor(a) responsável pelo conteúdo pedagógico do Ensino Fundamental da 5ª a 8ª série;

- um professor responsável pelo conteúdo pedagógico do Ensino Médio;

- dois professores/consultores para adequar os conteúdos e a linguagem pedagógica dos diferentes níveis;

- quatro professores, durante quatro meses, com expediente de seis horas, para dar atendimento ao telefone 0800 e responder às consultas sobre o conteúdo do material; e

- um *webdesigner* – técnico em Internet - para elaboração do site Vamos Contar, contendo informações sobre o projeto, apresentando atividades propostas nos Guias do Professor, além da mostra de parte do vídeo da D. Ruth Cardoso com o *jingle* do Censo 2000, cantado por um coral infantil.

A necessidade e a natureza do projeto levou o IBGE a propor uma parceria com o Ministério da Educação. Ocorreram reuniões entre representantes dos dois órgãos para apresentação e aprovação do projeto. Houve também o apoio do Ministério para a realização de um trabalho de consultoria, de forma a validar o conteúdo do projeto, em conformidade com os novos Parâmetros Curriculares Nacionais para o Ensino Fundamental e para o Ensino Médio. Os custos desse trabalho de consultoria foram arcados pelo Ministério de Educação, que fez as seguintes indicações para o trabalho envolvendo as atividades do ensino fundamental, a saber:

- um professor(a) responsável pelo conteúdo pedagógico de Geografia e História; e

- um professor(a) responsável pelo conteúdo pedagógico de Matemática e Estatística.

No momento seguinte, o Ministério da Educação, através da Secretaria de Ensino Médio, disponibilizou técnicos para validar o conteúdo pedagógico das atividades de Ensino Médio. Foram feitos relatórios de consultoria sobre as atividades propostas para os dois níveis de ensino: Ensino Fundamental – 1ª a 4ª série e 5ª a 8ª série – e Ensino Médio.

Outras parcerias foram buscadas, de forma a compartilhar custos e viabilizar a divulgação do projeto. Procurou-se receber apoio de instituições que vinham contribuindo com o movimento em curso no Brasil de buscar, por meio da educação, a formação de cidadãos capacitados visando com isso maior desenvolvimento para o país. O IBGE concretizou parcerias com as seguintes empresas e órgãos:

- TV Globo, através do programa Amigos da Escola, com o envolvimento de 24 mil escolas do Ensino Fundamental e elaboração e envio de folhetos explicativos sobre o censo e sobre o projeto Vamos Contar!;

- Coordenadoria Nacional para Integração da Pessoa Portadora de Deficiência – CORDE, da Secretaria Nacional dos Direitos Humanos doMinistério da Justiça; e

- TV Futura, através a criação e veiculação de peças promocionais, divulgando o projeto na televisão durante os meses de novembro e dezembro de 2000.

Obteve-se, ainda, a colaboração das seguintes instituições:

- Empresa Municipal MultiRio, com a criação e veiculação de peças promocionais do projeto durante a programação em canais de televisão; e

- Conselho Comunidade Solidária, através da pessoa da antropóloga Dra. Ruth Cardoso, presidente daquele conselho, que gravou peça promocional do projeto, veiculada em todas emissoras de televisão, no horário gratuito.

4.2.6 Empresas contratadas

A implementação do projeto, exigiu a contratação de empresas para as seguintes atividades:

- fornecimento de papel especial para confecção dos mapas, formato 1,10 x 0,9 m. A proposta era fazer uso de um tipo de papel que fosse resistente a manchas, rasgos, deformação, etc., já que o projeto previa a permanência dos mapas em cada sala de aula do país; no entanto, a opção foi pela compra e uso do papel comum com gramatura um pouco mais resistente;

- elaboração da fita de vídeo do projeto. Com duração de, aproximadamente, 10 minutos, o vídeo elaborado teve como finalidade apresentar, de forma sucinta, o projeto, destacando seu objetivo, o material envolvido e o público a que se destinava;

- impressão e envio de *folder* explicativo às escolas integrantes do projeto, em forma de mala direta, de modo a informarem os seus diretores sobre a realização do projeto durante o ano 2000 e solicitar seu apoio e aguardo do material durante o primeiro semestre do ano; o envio do *folder* aconteceu no final do ano de 1999;

- elaboração de três fitas de vídeo, com duração aproximada de 20 minutos cada, para os níveis de ensino apresentados, especificando o conteúdo pedagógico, com a finalidade de explicar e exemplificar as diferentes atividades sugeridas nos Guias do Professor;

- confecção de 220 000 caixas, de dois tamanhos diferentes, para acomodação dos kits com material didático;

- confecção de 60 000 envelopes para montagem dos kits;

- confecção de 750 000 embalagens (canudos) para 1 milhão de mapas produzidos; e

- fornecimento de mão-de-obra complementar para acondicionamento e postagem de todo o material do projeto; para envio desse material foi utilizado o contrato de porte pago já existente com a Empresa Brasileira de Correios e Telégrafos – ECT.

4.2.7 Teste

A partir de 1º de agosto de 1999, o IBGE realizou, nos municípios de Marília (SP) e Bonito (PA), o Censo Experimental, um ensaio com vistas ao Censo 2000. A operação teve como objetivo simular todas as etapas da realização do Censo, desde o planejamento, o recrutamento e seleção do pessoal, a coleta das informações, passando pela entrada de dados, chegando até a sua apuração. Na ocasião, foi também testado o projeto Vamos Contar!, em escolas

dos dois municípios, com a aplicação de questionários e realização de entrevistas – ver mais detalhes no Capítulo 5.

4.2.8 Implementação

Em 1º de novembro de 2000, deu-se início à postagem do material às escolas previstas do projeto, e para isso, contou-se com um cadastro obtido junto ao Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP; o atraso no envio do material foi motivado pela dificuldade de impressão das peças do projeto.

Essa etapa do projeto foi coordenada pela gráfica do IBGE, tendo sido necessário criar equipes de trabalho para gerenciar e auxiliar no acondicionamento do material. Foi previsto em contrato que a empresa realizasse o trabalho em cerca de dois meses; levou-se, aproximadamente, cinco meses para acondicionar 85% dos mapas.

O envio do material obedeceu ao seguinte plano de remessa:

- primeira etapa: 2 000 maiores escolas do país, contando com o auxílio das Unidades Regionais para distribuição do material;

- segunda etapa: 24 000 escolas do Programa Amigos da Escola, de forma a atender o convênio firmado entre o IBGE e a TV Globo; e

- terceira etapa: restante das quase 165 000 escolas; estima-se que 80% desse material foi entregue no prazo de quatro meses, com início em janeiro de 2001.

Apesar do atraso no envio do material, o Vamos Contar! foi uma iniciativa bem sucedida. A idéia inicial de mobilizar, previamente, alunos, professores e a comunidade escolar, para a realização do Censo 2000, não foi concretizada, mas acabou provocando uma mudança no rumo do projeto.

O foco principal do Vamos Contar! foi redirecionado para o trabalho com o material didático em sala de aula, de modo que os alunos pudessem dedicar-se mais ao estudo da estatística, aos trabalhos com gráficos e mapas, com o acréscimo das questões do Censo, explicando o mecanismo da operação e a importância da participação da sociedade. Essa mudança de foco permitiu, ainda, dar mais destaque aos resultados do Censo 2000, sua importância e aplicabilidade na solução dos problemas da comunidade.

O uso dos dados para conhecimento da realidade foi apresentado às crianças e jovens que participaram do projeto, dando a idéia do caminho percorrido para se obter uma determinada informação para uma localidade. Além dos exercícios práticos, de matemática, estudos sociais e outras disciplinas, os alunos tiveram a oportunidade de entender melhor a relação entre a sociedade, a informação e a melhoria de condições de vida.

4.3 Comissões Censitárias Municipais

As Comissões Censitárias Municipais – CCMs tiveram como objetivo buscar junto à sociedade organizada, em nível municipal, apoio e parcerias visando dar transparência ao Censo e obter facilidades para sua realização. Deviam, em nome da comunidade, acompanhar o desenvolvimento dos tra-

balhos da pesquisa, obedecendo as orientações definidas pelo IBGE, resguardando o sigilo das informações, contribuindo para a obtenção de informações fidedignas.

As CCMs foram estruturadas conforme a Resolução do Conselho Diretor – RC nº 0016, de 9 de dezembro de 1999, e o documento *Regulamentação das comissões censitárias municipais para o censo demográfico 2000* (2000).

No Censo 2000, as comissões municipais seguiram o mesmo modelo adotado em censos anteriores, com ajustes que visaram ao aprimoramento e padronização dos documentos utilizados e ao controle dos registros e da realização das reuniões. Além das orientações, contidas na regulamentação acima referida, a respeito da constituição e funcionamento das comissões e da organização das reuniões, foram também desenvolvidos modelos de atas e produtos de apoio, como folheto explicativo e filme que destacavam aspectos relevantes e a importância do funcionamento das comissões.

Idealizadas para terem um mínimo de cinco membros, foram presididas por representantes do IBGE, prevendo:

- dois representantes do Poder Executivo Municipal; o Prefeito ou seu representante e, preferencialmente o Secretário da Secretaria de Educação do Município, visando o envolvimento com o projeto Vamos Contar! ;

- um representante do Poder Legislativo, um do Poder Judiciário; e

- também representantes de universidades, instituições estaduais ou federais, líderes comunitários, sindicatos, federações, clubes, lojas maçônicas, empresas privadas, entre outros.

As CCMs tiveram como atribuições:

- colaborar na instalação do posto de coleta e auxiliar na disponibilização da infra-estrutura necessária à coleta de dados;

- analisar os mapas para fins estatísticos elaborados pelo IBGE para orientar as equipes de campo e a divulgação de resultados;– promover a divulgação do Censo junto à comunidade, buscando o envolvimento de pessoas que pudessem colaborar para a realização dos trabalhos censitários;

- acompanhar a evolução da coleta através da análise dos relatórios do Sistema de Indicadores Gerenciais da Coleta; e

- apresentar para o IBGE o nome e endereço das pessoas que não tivessem sido recenseadas.

No período de outubro de 1999 a março de 2001, foram instaladas 5 538 CCMs, gerando um total de 16 772 atas de reuniões, encaminhadas até dezembro de 2001 à Coordenação das Comissões Censitárias Municipais, instituída no Centro de Documentação e Disseminação de Informações do IBGE, assim distribuídas.

Tabela 4.1 - Atas das reuniões das Comissões Censitárias Municipais, segundo o tipo - 2000

Tipo	Quantidade
Atas de instalação	5 538
Atas de encerramento	5 533
Atas da 1ª reunião ordinária	3 545
Atas da 2ª reunião ordinária	1 534
Atas da 3ª reunião ordinária	523
Atas da 4ª reunião ordinária	96
Atas da 5ª reunião ordinária	2
Ata da 6ª reunião ordinária	1

Fonte: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos.

As atas foram disponibilizadas para consulta, no canal Censo 2000 do portal do IBGE na Internet, possibilitando conhecimento dos assuntos tratados nas CCMs.

4.4 Promoção local e campanha publicitária nacional

Dois pontos foram fundamentais para guiar a promoção e a publicidade do Censo 2000: sensibilizar a população para receber bem o Recenseador e mostrar a todos a importância de se responder corretamente ao questionário. Com base nesses dois objetivos, foram desenvolvidas ações de comunicação, tanto em nível local, aproveitando parcerias e oportunidades de cada lugar, como em nível nacional, com uma campanha veiculada pelos meios de comunicação de massa, para levar o Censo 2000 ao conhecimento do maior número de pessoas em todo o País.

4.4.1 Promoção local

As unidades do IBGE, nas capitais dos estados e também em outros municípios, desenvolveram ações locais de comunicação, que contribuíram para a divulgação do Censo 2000, tendo em vista a realização da coleta de dados e para o reforço da campanha publicitária nacional.

Cada unidade recebeu, do Centro de Documentação e Disseminação de Informações, um manual com explicações detalhadas sobre o uso da logomarca do Censo 2000, em diversos tipos de peças promocionais – camisetas, bonés, etc. – além de algumas idéias para iniciativas a serem aplicadas de forma a divulgar o Censo 2000. Entre as ações desenvolvidas localmente, pode-se citar:

- veiculação gratuita de *spots* – mensagem específica para ser veiculada em rádio, podendo ter trilha sonora ou não - e comerciais nas emissoras locais de rádio e televisão;

- parceria com instituições e empresas para produção de peças promocionais – bonés, sacolas de supermercado, camisetas, etc. – com a logomarca do Censo 2000;

- divulgação de mensagens sobre o censo, através de carros de som;

- veiculação de mensagens de reforço sobre a importância do censo, em contas de luz, água, telefone e gás;

- colocação de *banners* – cartaz de grande dimensão, produzido em material resistente - em *shopping centers*;

- exibição de mensagens nos placares eletrônicos de estádios de futebol; e

- divulgação, através de companhias aéreas, cujos comissários de bordo utilizaram o *pin* (broche) do Censo 2000.

4.4.2 Campanha publicitária nacional

Para a elaboração da campanha publicitária nacional de divulgação do Censo 2000 foi contratada uma empresa especializada em serviços de publicidade, através de processo licitatório na modalidade concorrência pública, iniciado em 13 dezembro de 1999, conforme publicação do edital no Diário Oficial da União.

Além das regras da licitação, o edital, para as agências concorrentes, apresentou também um *briefing*, com orientações e informações referentes ao IBGE e ao Censo, destacando os seguintes objetivos da campanha publicitária: mobilizar toda a sociedade brasileira para a importância de receber bem o Recenseador do IBGE e de responder corretamente ao questionário do Censo 2000.

Foi criada uma Comissão Especial de Licitação, formada por membros do IBGE e da Secretaria de Comunicação de Governo da Presidência da República, para acompanhar o processo licitatório, que teve três etapas básicas: exame dos documentos de habilitação das empresas concorrentes, julgamento da proposta técnica e valoração da proposta de preço.

Dentre as dez agências de publicidade que participaram da concorrência, a empresa vencedora foi a Standard, Ogilvy & Mather, que apresentou melhor técnica e melhor preço, conforme estabelecido no edital. O resultado final foi publicado no Diário Oficial da União de 13 de março de 2000.

A estratégia de comunicação construída pela agência vencedora tinha os objetivos de: informar, educar, motivar e mobilizar a população brasileira para o Censo 2000. O primeiro passo era informar, ao mesmo tempo educando e motivando, chamando a atenção para as vantagens do Censo. A mobilização surgiu na fase final, conseqüência desse trabalho de base. O conteúdo estratégico central da comunicação foi a expressão direta e emocional do grande objetivo da pesquisa: o Censo fez perguntas cujas respostas vão melhorar o futuro do país e de cada um dos seus cidadãos. Em torno deste conteúdo, e a partir dele, foram transmitidos, absorvidos e aceitos pelo público-alvo os demais objetivos: a necessidade de realização do Censo; a necessidade da participação; e a importância de responder corretamente. E ainda aqueles objetivos ligados à operação propriamente dita: a confiança no IBGE e na organização do Censo; a certeza da confidencialidade; e a tranqüilidade quanto à segurança pessoal e do domicílio.

O trabalho da agência de publicidade compreendeu as seguintes tarefas:

- estudo, a concepção, a execução e a distribuição de campanhas e peças publicitárias;

- assessoramento e apoio na execução de ações de promoção e patrocínio;

- assessoria de imprensa e relações públicas; e

- desenvolvimento de pesquisas e demais serviços relacionados com a comunicação do Censo 2000.

A seguir, é apresentado o quadro que lista os meios de comunicação de massa utilizados e inserções realizadas, detalhando a parte de mídia da campanha.

Quadro 4.1 - Meios de comunicação de massa, campanhas de mídia, utilizados na publicidade do Censo 2000

Meios de comunicação	Inserções
Rádio	25 emissoras com 3 059 inserções em agosto/2000 25 emissoras com 2 079 inserções em setembro/2000 Foram veiculados quatro comerciais diferentes Formatos: 30" e 60"
Revista	Seis revistas nos meses de junho a agosto/2000 Foram publicados quatro anúncios diferentes Formato: página dupla, quatro cores.
Jornal	22 jornais em dez praças na primeira semana de agosto/2000 Formato: meia página, quatro cores
Outdoor	780 *outdoors* em 27 praças na primeira quinzena de agosto/2000 780 *outdoors* em 27 praças na primeira quinzena de setembro/2000
Busdoor	846 adesivos em 24 praças em agosto/2000 60 adesivos em uma praça em setembro/2000
Painel eletrônico	17 painéis em dez praças em agosto/2000 com 360 inserções de 10 segundos por dia, em cada praça
TV aberta	Dez emissoras nos meses de julho a setembro Foram veiculados quatro comerciais diferentes Formatos: 30" e 60"
TV a cabo	Quatro emissoras nos meses de julho a setembro Foram veiculados quatro comerciais diferentes Formatos: 30" e 60"
Merchandising	*Merchandising* nos programas "Domingão do Faustão" e "Domingo Legal", com comercial chamado, para mercado nacional, com duração de 60" *Merchandising* com três ações na novela "Laços de Família", da TV Globo, em horário nobre
Internet	Colocação de *banners* em quatro portais durante agosto e setembro/2000, com 17 750 000 *page views* (cliques dos internautas)

Fonte: IBGE, Centro de Documentação e Disseminação de Informações.

Já o quadro abaixo detalha o item produção, com a listagem das peças impressas, dos brindes, dos comerciais, e de todo o material produzido veiculado.

Quadro 4.2 - Quantidades de peças produzidas para a campanha publicitária do Censo 2000

Peça	Quantidades e outras informações
Cartaz	Produção de seis tipos diferentes de cartazes, sendo impressos 600 000 (100 000 cartazes de cada tipo).
Kit Vamos Contar! Para escolas	Impressão de 1 000 000 de etiquetas de endereçamento; 2 000 000 de mapas; 1 550 000 cartas ao professor; 200 000 cartas aos membros do Conselho Escola-comunidade; 200 000 cartas ao diretor; 1 550 000 fichas de avaliação; 2 300 000 folhetos; 820 000 guias para professor de 1ª a 4ª série do Ensino Fundamental; 500 000 guias para professor da 5ª a 8ª série do Ensino Fundamental; e 200 000 guias para professor do Ensino Médio.
Vídeo Vamos Contar! – treinamento para professores	Produção de vídeo específico para professores, com aplicação de todos os guias do Vamos Contar! e apresentação geral do projeto. Foram feitas 5 600 cópias para distribuir às unidades do IBGE e atender escolas de todos os municípios.
Comercial para TV Vamos Contar!	Produção de comercial com D. Ruth Cardoso sobre o Vamos Contar!
Site Vamos Contar!	Disponibilização de informações sobre o Vamos Contar! na Internet.
Pin	Confecção de 50 000 broches com logomarca do Censo 2000.
Mouse pad	Produção de 10 000 *mouse pads* com logomarca do Censo 2000, *e-mail* e endereço na Internet.
Carta ao síndico	Impressão de 1 000 000 de cartas, 1 000 000 de envelopes.
Bloco do recenseador	Impressão de 5 000 000 de lâminas, para recados aos moradores não encontrados nos domicílios.

Fonte: IBGE, Centro de Documentação e Disseminação de Informações.

O valor do contrato de prestação de serviços de publicidade foi de R$ 17.000.000.000,00, válido por um período de 12 meses. Nos três anos seguintes, em função da divulgação dos resultados do Censo 2000 e do lançamento dos produtos contendo esses resultados, o contrato com a agência de publicidade foi renovado ano a ano, com os seguintes valores: em 2001, R$ 1.500.000,00; em 2002, R$ 600.000,00; e em 2003, R$ 350.000,00; a partir de 2001, não houve publicidade através da mídia.

A partir de 2001, a verba de publicidade foi utilizada na produção de peças impressas, eventos de divulgação dos resultados, ações especiais de comunicação, produção de vídeo, assessoria de imprensa e outros serviços que promovessem os produtos lançados com resultados do Censo.

O IBGE informou à Secretaria de Comunicação de Governo da Presidência da República, responsável pela aprovação da verba de publicidade, que o contrato não seria renovado para o ano de 2004, já que os trabalhos de divulgação se encerrariam com o lançamento das últimas publicações do Censo 2000.

4.5 Atendimento telefônico

De janeiro a dezembro de 2000, a linha 0800-218181 do IBGE passou a atender a chamadas referentes às atividades do Censo 2000, em um total de, aproximadamente, 247 mil; desse número, cerca de 50% foram ligações relativas aos processos de seleção de pessoal temporário. A estrutura do atendimento telefônico contou com até 31 atendentes – nos períodos de pico, distribuídos em quatro turnos, funcionando de 7h às 22h, de segunda a domingo.

O 0800 foi um serviço amplamente divulgado durante a campanha promocional e publicitária do Censo 2000, vindo a constituir-se em importante canal de comunicação entre o IBGE e a população, durante o período de coleta das informações – ver Capítulo 9.

Na fase de mobilização, anterior à coleta do Censo, o que foi objeto de consulta através do atendimento telefônico, estava relacionado aos seguintes assuntos:

- inscrições, resultados e outras informações sobre os processos seletivos para contratação de pessoal temporário;

- informações sobre contratações;

- operação censitária; e

- Projeto Vamos Contar!.

A tabela abaixo mostra o impacto da preparação para o Censo 2000 no atendimento telefônico do IBGE, comparando o serviço nos meses que antecederam à pesquisa no ano 2000, com os correspondentes ao ano anterior. Vale destacar que, em 2000, o 0800 recebeu 247 429 ligações, enquanto durante o ano de 1999 foram 11 216 chamadas.

Tabela 4.2 - Ligações recebidas na linha 0800 do IBGE, por período de atendimento - janeiro 1999/julho 2000

Ano	Ligações recebidas							
	Total	Período de atendimento						
		Janeiro	Fevereiro	Março	Abril	Maio	Junho	Julho
1999	6 230	891	655	870	833	1 322	665	994
2000	144 139	4 622	22 659	21 400	24 769	31 256	20 457	18 976

Fonte: IBGE, Centro de Documentação e Disseminação de Informações, Departamento de Atendimento Integrado.

4.6 Censo 2000 na Internet

O canal Censo 2000 foi desenvolvido com o objetivo de estreitar as relações com a sociedade. A página entrou no ar em 1997, integrando o portal do IBGE. Pelo endereço http://www.ibge.gov.br/censo, foi possível conhecer as ações desenvolvidas para a mobilização da população, um pouco da história do Censo, a legislação, a importância do Censo, os motivos para respondê-lo, etc. Divulgaram-se as comissões constituídas, os parceiros externos, os questionários, os textos técnicos e o cronograma de trabalho. As principais informações do Censo tiveram divulgação simultânea na página do Censo e na imprensa.

A página recebia atualizações à medida que havia novas informações sobre a operação; seu menu principal apresentava os seguintes itens e respectivos *links*:

História

- a história do Censo no Brasil;
- a história do Censo no mundo;

População

- crescimento da população brasileira;
- Censo Demográfico 1991;
- Contagem da População 1996;
- Censo Demográfico 2000;

Censo em Foco

- Censo Demográfico e políticas públicas;
- a obrigatoriedade e o sigilo;
- a periodicidade do Censo;
- a escolha da data de realização;
- os números do Censo;

Organização

- legislação;
- orçamento;
- base operacional;
- a coleta;
- Censo Experimental 1999
- textos técnicos;
- cronograma;
- plano de divulgação;

Mobilização

- ações preparatórias;
- parceiros externos;
- a propaganda do Censo 2000;

Comissões

- Comissões Censitárias Municipais; e
- CPO e Comissão Consultiva.

Outros *links* compunham a página, a saber: a importância do Censo 2000; quem foi recenseado; os motivos para responder ao Censo; as inovações do Censo 2000; divulgação dos resultados; Mercosul; Projeto Vamos Contar!; revista do censo; FAQ/webmaster; e outros *links*.

Os internautas tinham, ainda, disponíveis um serviço de respostas às perguntas mais freqüentes e um endereço de correio eletrônico – censo2000@ibge.gov.br – para enviar perguntas e sugestões – ver Capítulo 9. A página do Censo 2000 também divulgou o número do atendimento telefônico do IBGE 0800-218181.

4.7 Revista Vou te Contar

Assim como ocorreu no Censo de 1991, quando foi editado o jornal *Censo*, em 2000, elegeu-se um meio de comunicação das atividades de planejamento, coleta, apuração e divulgação. A *Vou te contar: revista do censo 2000* foi criada com a finalidade de funcionar como um veículo de comunicação e de integração entre as áreas envolvidas na realização do Censo 2000 e divulgar as ações referentes a essa operação.

A revista teve uma tiragem de 6 000 exemplares por edição e, de 2000 a 2003, foram editados 12 números. A primeira edição apresentava oito seções. Já na segunda edição, devido à boa repercussão do primeiro número, a revista ganhou mais três seções e quase dobrou o número de páginas.

Em termos técnicos, a *Vou te contar* teve as seguintes características:

- formato 21 x 27cm (fechado) e 42 x 27cm (aberto);

- impressão 4/4 cores (miolo colorido) e 4/1 (capa colorida e anúncios internos em preto e branco);

- acabamento com grampo; e

- número de páginas variando de 32 a 44, incluindo capa - com exceção da primeira que teve 20 páginas.

A seguir estão destacadas as seções da revista e o tipo de conteúdo que apresentavam:

- Editorial – a palavra do presidente do IBGE, com ênfase na etapa do Censo na época de cada edição, apresentando, ainda, uma visão geral dos temas que seriam abordados;
- Conta-gotas – notas referentes aos censos de outros países: curiosidades, inovações, resultados, etc.;
- Matéria de capa – textos e entrevistas referentes ao assunto da capa da edição;
- Espaço aberto – seção de formato mais livre, que tinha sempre um tema central, podendo ser apresentado através de entrevista – pergunta e resposta - de artigo ou texto realizado a partir de pesquisas e entrevistas;
- Gente contando gente – entrevista, sempre no formato pergunta e resposta, com personalidades que tivessem atuação destacada no Censo 2000, podendo ser do próprio IBGE ou não;
- Nos estados – cobertura da participação das unidades do IBGE, em todo o Brasil, no Censo 2000;
- Reportagem – espaço destinado à cobertura de assuntos mais extensos, como atividades envolvendo áreas internas, grandes projetos, entre outros;

- Registro – como o próprio nome diz, era o registro, geralmente através de textos curtos, de novos produtos a serem lançados, eventos e outras iniciativas dentro do Censo 2000;
- Censo em foco – textos sobre ações e eventos relacionados ao Censo 2000, tratados mais pela sua representatividade do que pelo momento em que aconteceram;
- Atualidades – as novidades do Censo 2000, recém-lançadas ou ainda por lançar. Foram destaque nessa seção, por exemplo, o vídeo feito pelo cineasta Zelito Viana a partir de dados do Censo e o cronograma de divulgação dos resultados do Censo 2000; e
- Ponto-de-vista – em geral, continha um artigo de técnico do IBGE ou externo ao instituto.

Os textos eram acompanhados de ilustrações e fotos. A revista continha, ainda, anúncios de produtos do IBGE na segunda, terceira e quarta capas. Durante todas as suas edições, manteve um alto padrão gráfico. Devido ao grande volume de material produzido na gráfica do IBGE, para a realização do Censo 2000, alguns números da revista foram impressos por gráfica externa, sempre mantendo o padrão gráfico. A revista *Vou te contar* tinha, como destinatários, as unidades do IBGE no Brasil e institutos de estatística no exterior.

O CD-ROM, anexo a essa publicação contém os arquivos pdf das edições da revista *Vou te contar*.

Referências

IBGE. Resolução do Conselho Diretor- CD nº 0016, de 9 de dezembro de 1999. Determina aos titulares das Divisões de Pesquisas – DIPEQ que providenciem a constituição de uma Comissão Censitária Municipal, em cada município de sua jurisdição, com o objetivo de acompanhar o desenvolvimento da operação de coleta do Censo Demográfico 2000. Rio de Janeiro, 1999. Disponível em: <http://w3.aplicacao.de.ibge.gov.br/Siad/SIAD.nsf>. Acesso em: out. 2003.

REGULAMENTAÇÃO das comissões censitárias municipais para o censo demográfico 2000. Rio de Janeiro: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, 2000. 20 p.

RELATÓRIO de avaliação da etapa de coleta do censo 2000. Rio de Janeiro: IBGE, Diretoria Executiva, 2001. 22 p.

VOU TE CONTAR: revista do Censo 2000. Rio de Janeiro: IBGE, 2000- . Bimestral.

5 Provas Piloto e Censo Experimental

Os testes prévios à realização do Censo Demográfico 2000 constituíram-se em elementos fundamentais para o seu sucesso.

As provas piloto revelaram-se proveitosas para o esclarecimento e para o aperfeiçoamento das intenções de investigação ainda incompletas ou pouco elaboradas.

O avanço assim obtido, tornou possível realizar no Censo Experimental, uma simulação completa da operação censitária, constituindo uma proposição de pesquisa bastante avançada, no que tange à conformidade dos conceitos, aos questionários, ao treinamento das equipes, ao acompanhamento e controle da coleta e ao material utilizado nas entrevistas – lápis, borracha, prancheta etc.

5.1 Primeira Prova Piloto

A Primeira Prova Piloto do Censo 2000, realizada em dezembro de 1997, teve como objetivos principais:

- testar três modelos de questionários: o primeiro, mais reduzido, denominado Questionário Básico e os outros, mais extensos, identificados, respectivamente, como Questionário da Amostra 1 e Questionário da Amostra 2; e

- avaliar um processo automatizado de entrada de dados, utilizando a tecnologia de digitalização através de *scanner*, e de reconhecimento ótico de caracteres.

Foram testadas, também, diferentes alternativas de investigação para um mesmo tema, seqüências distintas de investigação de quesitos e blocos de quesitos. Além disso, a prova piloto avaliou os seguintes itens:

- fluxo de trabalho da pesquisa, incluindo distribuição do material, coleta das informações, acompanhamento e controle da coleta, digitalização dos questionários, reconhecimento óptico de caracteres e transmissão de resultados;

- tempo de preenchimento dos questionários;

- cor e desenho dos questionários;

- utilização de setas e comandos para a indicação da seqüência de preenchimento dos quesitos;

- velocidade do *scanner* no processo de digitalização dos questionários;

- adequabilidade do tipo do papel utilizado e da disposição dos quesitos dos questionários ao processo de reconhecimento óptico de caracteres; e

- parâmetros da qualidade do reconhecimento para cada um dos tipos de campos – de marca, numérico e alfanumérico – dos questionários.

A seguir, serão apresentados, de forma geral, os conteúdos dos três questionários testados. Informações mais detalhadas podem ser obtidas pela consulta a esses instrumentos de coleta, apresentados em anexo.

Características do Questionário Básico

O Questionário Básico investigou, no bloco Características da Habitação, a espécie, o tipo do domicílio particular permanente, o tipo do domicílio coletivo e a existência de iluminação elétrica.

No bloco Características dos Moradores, foram pesquisados o sexo, a relação com a pessoa de referência do domicílio, o mês e o ano de nascimento, a idade presumida e a cor ou raça. Havia, também, no Questionário Básico, espaço para a lista de moradores do domicílio.

Características dos Questionários da Amostra 1 e da Amostra 2

Os Questionários da Amostra 1 e da Amostra 2 – também tratados neste texto como questionários 1 e 2 – apresentaram, entre si, algumas diferenças significativas quanto à forma de apresentação e ao conteúdo dos quesitos.

O Questionário da Amostra 1 pesquisou, no bloco Características da Habitação, todas as informações do Questionário Básico e mais as seguintes características:

- material das paredes externas, do piso e da cobertura;

- número de cômodos e de cômodos servindo de dormitório;

- condição de ocupação do domicílio e do terreno onde se localizava;

- forma de abastecimento de água e de sua canalização no domicílio;

- destino do lixo, existência de banheiro e o tipo de escoadouro utilizado; e

- existência de linha telefônica instalada ou aparelho celular, rádio, televisão e geladeira.

No Questionário da Amostra 2, não foi investigado o material das paredes externas do domicílio. No entanto, além de todas as demais informações citadas anteriormente, esse instrumento de coleta pesquisou os seguintes pontos:

- tipo da parede externa;

- existência de revestimento e pintura das paredes internas;

- existência de banheiro com vaso sanitário, chuveiro e pia;

- existência de cozinha com fogão, geladeira e pia; e

- existência de automóvel particular, televisão em cores e preto e branco, videocassete, *freezer*, máquina de lavar roupa, forno de microondas e microcomputador.

No Questionário da Amostra 2, a investigação dos bens duráveis, além de mais extensa, incluiu também a quantidade de cada um dos bens.

No que se refere ao bloco Características dos Moradores, os Questionários da Amostra 1 e da Amostra 2, além de investigarem todas as informações do Questionário Básico, pesquisaram também:

- a origem e a relação com a pessoa de referência da família;

- a religião e a existência de mãe e pai vivos;

- se participava ou era beneficiário de plano de saúde privado; e

- diversos aspectos sobre os temas deficiência, migração, educação, nupcialidade, fecundidade e trabalho e rendimento.

Nos dois questionários da amostra, investigou-se, pela primeira vez, no bloco sobre nupcialidade, "a união entre pessoas do mesmo sexo". No Questionário da Amostra 1, o quesito sobre a origem da pessoa antecedia o quesito "cor ou raça", enquanto no da Amostra 2 acontecia o contrário.

O bloco sobre deficiência, no Questionário da Amostra 1, investigava se a pessoa tinha algum tipo de "deficiência - física ou mental" – enquanto no Questionário da Amostra 2, buscava-se, também, investigar a "auto-avaliação sobre a capacidade da pessoa enxergar, ouvir, caminhar e aprender".

O bloco sobre trabalho e rendimento, embora não tivesse sofrido alterações significativas em seu conteúdo, entre os dois questionários da amostra, passou por consideráveis modificações na ordem seqüencial de seus quesitos, num e noutro modelo.

Uma outra diferença refere-se ao bloco de fecundidade que, no Questionário da Amostra 1 antecedia o bloco sobre trabalho e rendimento e passou a ser investigado por último, no Questionário da Amostra 2.

Âmbito da pesquisa

A Primeira Prova Piloto foi realizada nos Estados do Rio de Janeiro e Mato Grosso. Esses estados foram escolhidos levando-se em conta três fatores fundamentais: o custo da operação; a disponibilidade do corpo técnico do IBGE; e a representatividade das características de boa parte do território brasileiro, nas áreas urbana e rural.

No Rio de Janeiro, a única questão contrária ficou por conta da sua área rural, que não apresentava certas características de parte da Região Centro-Sul. No entanto, o Estado dispunha de uma rede de coleta bem estruturada e, por ser o Estado sede do IBGE, contribuía, decisivamente, para a redução dos custos operacionais. Os Municípios escolhidos foram Campos dos Goitacazes e Teresópolis.

A indicação do Estado de Mato Grosso ocorreu em função de apresentar características semelhantes às das Regiões Norte e Centro-Oeste. Foram selecionados para a pesquisa, ao norte do Estado, o Município de Sinop, que apresenta características amazônicas, e a sudoeste, o Município de Cáceres, representando a Região do Pantanal.

A coleta da Primeira Prova Piloto foi realizada em 18 setores de cada um dos estados escolhidos. Em cada município pesquisado, foram selecionados cinco setores urbanos e quatro rurais, escolhidos pelas Divisões de Pesquisas das Unidades Regionais do IBGE.

Treinamento

O treinamento, realizado no final de novembro de 1997, contou com a participação de 21 técnicos de cada Unidade da Federação envolvida na prova piloto. Para cada uma, foram indicados dois técnicos da DPE que atuaram como instrutores.

A capacitação das equipes teve como base as instruções do Manual do Recenseador que, além de tratar os conceitos e formas de investigação da pesquisa, ressaltou as inovações processadas na estrutura dos questionários, e os cuidados quanto às normas de preenchimento exigidas para o processo de digitalização e reconhecimento óptico de caracteres.

Trabalho de campo

Em cada Unidade da Federação, foram visitados 1 800 domicílios particulares. Para cada setor selecionado, foi utilizado um recenseador, técnico do quadro do IBGE, que entrevistou um total de 100 domicílios, considerando-se um período de 10 dias de coleta para os setores urbanos e de 15 dias para os setores rurais. Todo esse trabalho foi acompanhado pelo chefe da agência do IBGE no município, e por técnicos da sede do IBGE que atuaram como observadores durante a operação.

Orientando-se pela Caderneta do Setor e Folhas de Coleta utilizadas na Contagem da População 1996, o Recenseador iniciou o percurso do setor, partindo do mesmo ponto assinalado no mapa. Seguindo a ordem original do percurso estabelecido, a cada domicílio encontrado, foi aplicado um modelo do questionário. Portanto, os domicílios eventualmente omitidos na contagem ou construídos após aquela operação, não foram objeto de pesquisa na prova piloto. O domicílio ocupado durante a contagem e que estava vago ou fechado, na data da aplicação do questionário da prova piloto, foi substituído por outro, mantendo-se a mesma ordem de percurso observada na Folha de Coleta.

Não era objetivo dessa prova piloto testar os procedimentos de seleção da amostra. Optou-se por aplicar, para cada Questionário Básico, nove questionários dos demais modelos. O processo de seleção dos domicílios ocupados, para a aplicação dos três tipos de questionários, obedeceu à ordenação das unidades na Folha de Coleta e ao seguinte critério:

- para os primeiros 50 domicílios, foi aplicado o Questionários Básico em todos aqueles cujos números de ordem tivessem zero como último algarismo; nos demais domicílios, foi utilizado o Questionários da Amostra 1; e

- nos outros 50 domicílios, continuava-se com o mesmo critério para o uso do Questionário Básico e, aplicava-se, nos demais domicílios, o Questionário da Amostra 2.

Financiamento da operação

Os recursos previstos para o Censo no ano de 1997 cobriam, apenas, os custos da operação de campo da Primeira Prova Piloto. Desta forma, recorreu-se ao Fundo de População das Nações Unidas – FNUAP – com a apresentação do projeto "Implementação de Novas Tecnologias na Captura de Dados do Censo Demográfico do Ano 2000: Primeira Prova Piloto". Através deste projeto, garantiu-se que a entrada dos dados da pesquisa fosse feita através de *scanner*, utilizando o sistema de reconhecimento ótico de caracteres e contando com a consultoria de técnicos que realizaram esse trabalho no Uruguai. Além disso, os recursos financiaram a compra de um *scanner*, a contratação de pessoal para operação e outras despesas para a impressão dos questionários, incluindo o papel especial.

Avaliação da operação

A Primeira Prova Piloto foi avaliada, sob o ponto de vista do trabalho de coleta, através de relatórios individuais, preparados pelos recenseadores e observadores, consolidados no documento *Primeira prova piloto: relatório sobre a primeira prova piloto do censo demográfico 2000* (1997). Houve também uma avaliação no que toca ao processo automatizado de entrada de dados. A seguir, apresenta-se uma síntese dessas avaliações.

a) Trabalho de campo

Apresenta-se, a seguir, um conjunto de pontos, que resume o conteúdo do relatório consolidado.

Avaliação dos tempos médios de entrevista

Foram avaliados os tempos médios de entrevista para cada tipo de questionário aplicado por Unidade da Federação. Pôde-se observar que, entre os Questionários da Amostra 1 e 2, praticamente não houve diferença entre os tempos médios de entrevista por pessoa, a saber: 9,5 minutos e 9,3 minutos, respectivamente. Também, entre as Unidades da Federação, não houve diferença expressiva.

Verificou-se uma significativa diferença na área rural de Mato Grosso, que apresentou um tempo médio de 10,4 minutos, enquanto a área urbana apresentou uma média de 8,2 minutos. No Rio de Janeiro, praticamente não houve diferença entre as situações urbana - 8,3 minutos - e rural - 8,2 minutos.

Análise comparativa dos resultados entre os questionários

Os resultados da aplicação dos Questionário das Amostras 1 e 2 foram comparados quanto aos componentes de estrutura do domicílio – "parede", "piso" e "cobertura". Apesar de não se ter verificado diferença significativa entre os questionários, neste aspecto preferiu-se optar pelo Questionário da Amostra 1, onde o Recenseador classificava o domicilio somente de acordo com a observação do material utilizado na construção, enquanto, no outro, deveria classificá-lo como durável ou não durável, utilizando critérios mais subjetivos.

Também foram realizadas comparações com relação à existência de rádio, geladeira e televisão, onde não se detectou diferença significativa. Neste caso, resolveu-se optar pelo Questionário da Amostra 2 pois, além da posse, investigava, também, a quantidade desses bens.

Com relação à característica "cor ou raça", observou-se uma diferença mais expressiva entre as classificações parda no Questionário da Amostra 1 - 39,2% - e mestiça no Questionário da Amostra 2 - 33,8%.

Os quesitos de deficiência apresentaram-se de maneira bastante distinta nos dois questionários da amostra. Alguns quesitos como "surdez total" e "cegueira total", que permitiam comparação, mostrou melhor adequabilidade do questionário 2, pois os quesitos são mais detalhados; além disso, são investigados de maneira indireta, com varias categorias para cada deficiência, enquanto no questionário 1 só se identifica sua existência. Os demais quesitos, embora não fossem totalmente comparáveis, não apresentaram resultados muito distintos.

A variável "número de filhos tidos nascidos mortos" apresentou muitos problemas de inconsistência, causando restrições quanto ao número de observações válidas para análise, o que dificultou possíveis comparações entre os questionários. O mesmo aconteceu com os dados sobre ramo de atividade, para crianças com idade entre cinco e nove anos, que apresentaram freqüência muito baixa, dificultando possíveis comparações.

Os quesitos de mão-de-obra das pessoas com dez anos ou mais apresentaram resultados praticamente iguais com relação às pessoas ocupadas - 54% no questionário da amostra 1 e 53% no questionário da amostra 2. Já o percentual de pessoas não ocupadas no ano de referência apresentou diferença mais significativa - 5,2% no questionário da amostra 1 e 3,4% no questionário da amostra 2. Neste sentido, foi acertado adotar a ordem de investigação do questionário da amostra 2 para esses quesitos.

Também foram avaliadas informações sobres pessoas responsáveis pelo domicílio, número de moradores por domicílio, e o número de informantes por domicílio, não se encontrando diferenças significativas entre os dois questionários da amostra.

Assim, em relação ao teste de conteúdo de investigação, pode-se afirmar que a prova piloto atendeu às expectativas pois, para vários pontos investigados, foi apontada a necessidade de correções e melhorias.

Avaliação geral e propostas de alterações

No caso do Manual do Recenseador, julgou-se conveniente efetuar todas as modificações sugeridas, uma vez que detectaram-se vários problemas de aplicação de conceitos por parte dos Recenseadores.

Para os questionários, julgou-se necessário as seguintes alterações:

- modificar a cor de fundo referente às características da habitação no questionário da amostra;

- dar maior destaque às setas de comando;

- dar maior relevo aos títulos que destacam a população alvo de cada bloco temático;

- enriquecer, sempre que possível, o questionário com lembretes conceituais para os principais quesitos ou blocos de quesitos;

- evitar quesitos com opções de resposta que exijam perguntas adicionais, não impressas no questionário; e

- aprimorar/retificar os textos das perguntas que causaram dúvidas de entendimento para informantes e para recenseadores.

Em relação ao desempenho dos Recenseadores, com base nos resultados de campo, decidiu-se reforçar, em futuros treinamentos, a necessidade de se respeitar, rigorosamente, o texto das perguntas, bem como observar os períodos de referência.

b) Processo automatizado de entrada de dados

O desenho dos questionários foi, em geral, considerado fator positivo para a garantia de qualidade na marcação dos campos por parte dos Recenseadores. Por outro lado, a qualidade da impressão não atendeu aos padrões necessários, ocasionando alguns erros de preenchimento.

O desenho gráfico dos campos de marca e seu formato – pequenos retângulos – funcionaram bem. Também obteve-se muito bom resultado com a apresentação de quadrículas pré-assinaladas, com letras e números no topo de cada página, que pôde servir de modelo para o Recenseador. Quanto ao tamanho dos campos não quadriculados, destinados à escrita, foi considerada a possibilidade de redução, com a finalidade de diminuir o espaço, trazendo melhor apresentação ao questionário.

Alguns erros de impressão, concentrados em certas páginas, acarretaram uma forte concentração de erro de leitura de alguns campos de marca. A magnitude do erro, fora dos padrões para esses campos, indica claramente que a origem do erro foi externa ao processamento. Estes erros seriam completamente eliminados com uma melhor qualidade de impressão.

A experiência com o reconhecimento das informações investigadas através de perguntas abertas foi muito positiva e demonstrou a viabilidade do uso da tecnologia empregada. A qualidade geral do preenchimento foi aceitável, sobretudo considerando que se tratava de uma primeira experiência e contribuiu para o desempenho do processo. O nível de reconhecimento foi bom, alcançando uma porcentagem global de 72%. Viu-se também que, com o uso de dicionários, mesmo de qualidade média de elaboração, podia-se alcançar, rapidamente, níveis de reconhecimento mais altos, superando 80%.

No processo de reconhecimento dos campos numéricos foi obtido bom resultado, apesar do registro desse campo apresentar menor confiabilidade, uma vez que possui baixa capacidade de contextualização. Em média, os campos numéricos são os que requerem maior percentagem de verificação humana, em

especial nos casos em que a informação é crítica e não existam formas simples e diretas de controlá-la.

Os resultados gerais do reconhecimento automático destes campos alcançou cerca de 65% de reconhecimento direto, requerendo correção humana somente em 7,64% dos casos.

Algumas considerações sobre o resultado do processo de digitalização e reconhecimento são apresentadas a seguir:

- os tempos de processamento e verificação manual foram muito bons. Houve problemas apenas na etapa de digitalização, em razão da grande quantidade de páginas de cada questionário;

- o sistema mostrou poder funcionar com equipamento de pequeno porte, acoplado a cada *scanner* – em torno de três microcomputadores de potência média, em cada *scanner* com capacidade de processamento de 50 páginas por minuto; também demonstrou ser fácil a agregação ou retirada de equipamentos na rede;

- o nível de intervenção humana no processamento foi muito baixo. Ao longo de três dias de configuração e ajuste, foi obtido um nível médio geral de 75% de reconhecimento automático, dado que devia alcançar, provavelmente, 85% ou 90% com o aumento do volume de dados, o que permitiria um enriquecimento dos dicionários; e

- o sistema de processamento mostrou não requerer, para ajustes e otimização dos aspectos operacionais e contextuais do processamento, a presença de profissionais da área de informática, apoiando o trabalho cotidiano. Essa situação apontava para que o sistema pudesse funcionar de maneira descentralizada, situação em que a presença permanente de especialistas em informática é muito reduzida.

5.2 Segunda Prova Piloto

Dando continuidade ao programa de testes, com vistas à realização do Censo, iniciado com a Primeira Prova Piloto, foi levada a campo a Segunda Prova Piloto do Censo Demográfico 2000, durante o mês de agosto de 1998.

Em relação ao teste anterior, foi promovido um aumento no número de áreas pesquisadas, feitas a introdução e a avaliação de outros instrumentos de coleta, ampliado o número de domicílios visitados e posta em prática uma rotina de treinamento bem mais elaborada. Outra novidade foi o teste de distribuição prévia e autopreenchimento dos questionários.

5.2.1 Objetivo

De modo geral, o objetivo da Segunda Prova Piloto foi avaliar a operação de campo como um todo, bem como o sistema de acompanhamento informatizado da coleta.

Em paralelo, foi realizado o teste de distribuição prévia e autopreenchimento dos questionários, cujo objetivo foi avaliar a receptividade da população a esse método de coleta e a qualidade das respostas, identificando as

dificuldades de preenchimento e de compreensão das instruções. A seguir, são apresentados os itens que foram objeto de avaliação.

relativos à operação de campo:

- forma e conteúdo dos questionários;

- forma e conteúdo dos seguintes instrumentos auxiliares da coleta: Caderneta do Setor, Folha de Coleta, Folha de Domicílio Coletivo, Folha Resumo do Setor, Folha de Apuração Preliminar e Folha de Encaminhamento de Setores;

- conteúdo do Manual do Recenseador;

- procedimentos de supervisão;

- treinamento para Recenseadores;

- desempenho dos Recenseadores contratados;

- tempo de entrevista; e

- processo de apuração preliminar.

relativos ao ambiente de informática:

- sistema de comunicação e liberação de informações entre a agência de coleta e a sede da Unidade Regional e desta com o órgão central do IBGE;

- sistema de acompanhamento e avaliação da coleta, de obtenção dos resultados preliminares e de recepção do material; e

- recursos de *hardware*.

5.2.2 Âmbito da pesquisa e características investigadas

A Segunda Prova Piloto foi realizada nos seguintes municípios:

- coleta por entrevista – Belo Horizonte e Governador Valadares, em Minas Gerais; Goiânia e Rio Verde, em Goiás; João Pessoa e Campina Grande, na Paraíba; e

- coleta por distribuição prévia e autopreenchimento – Rio de Janeiro e São Paulo.

Foram trabalhados, ao todo, 90 setores, sendo 20 em Belo Horizonte e Goiânia, 10 em Governador Valadares e Rio Verde, e 15 em João Pessoa e Campina Grande; nos municípios das capitais, todos os setores eram urbanos; nos demais municípios, foram pesquisados cinco setores rurais, sendo os demais urbanos.

Por outro lado, para o teste de distribuição prévia, utilizou-se, aproximadamente, 20 000 domicílios – divididos, igualmente, entre Rio de Janeiro e São Paulo – selecionados em áreas que congregavam a população com perfil mais adequado ao autopreenchimento dos questionários.

Características investigadas

Durante o processo de definição do conteúdo da investigação do teste, foram realizadas, no primeiro trimestre de 1998, reuniões temáticas com técnicos do IBGE envolvidos nessa etapa do planejamento. Realizaram-se, ainda, a análise e a sistematização das propostas de conteúdo, bem como as formas de investigação apresentadas pelos usuários externos.

Na Segunda Prova Piloto, foram aplicados dois questionários distintos, apresentados em anexo: o Questionário Básico, com um conjunto mais reduzido de quesitos e outro, mais extenso, o Questionário da Amostra.

O Questionário Básico investigou:

- no bloco Características do Domicílio – quesitos sobre espécie do domicílio, tipo do domicílio particular, tipo do domicílio coletivo, condição de ocupação do domicílio e do terreno onde se localizava, forma de abastecimento e canalização de água, número de banheiros, existência de instalação sanitária e tipo de uso, tipo de escoadouro da instalação sanitária e o destino do lixo;

- a lista de moradores do domicílio; e

- no bloco Características dos Moradores – quesitos básicos como sexo, idade, relação com a pessoa responsável pelo domicílio, e se sabe ler e escrever; somente para a pessoa responsável ou individual, em domicílio coletivo, investigou-se, ainda, a última série concluída com aprovação, o grau correspondente, bem como o rendimento bruto do mês de julho de 1998.

O Questionário da Amostra investigou as seguintes características do domicílio, além daquelas incluídas no Questionário Básico:

- material das paredes externas, do piso e da cobertura, revestimento interno das paredes, número total de cômodos e daqueles servindo de dormitório; existência de iluminação elétrica, linha telefônica, rádio, máquina de lavar, forno microondas; e

- número de videocassetes, televisores, geladeiras, microcomputadores, automóveis particulares e cômodos servidos por ar condicionado.

Em relação às características das pessoas, a investigação foi a mesma do Questionário Básico, com os seguintes acréscimos:

- quesitos relação com a pessoa responsável pela família, número da família, cor ou raça, origem e religião;

- conjuntos de quesitos sobre deficiência física, mental e auto-avaliação de ouvir, enxergar e caminhar, migração, escolaridade, estado civil e conjugal, fecundidade e mortalidade, mão-de-obra e rendimento; e

- lista de pessoas falecidas no último ano, investigando o sexo, a idade e o mês e ano de falecimento.

5.2.3 Treinamento e coleta dos dados

O treinamento dos Recenseadores para a coleta da Segunda Prova Piloto contou com um projeto desenvolvido pela COC, em conjunto com os técnicos da DPE. Esse projeto objetivou padronizar a utilização dos recursos instrucio-

nais, o uso das técnicas de ensino/aprendizagem e a instrução dos conceitos e procedimentos, visando levar o treinando a uma melhor assimilação das instruções, com vistas à obtenção de uma coleta de qualidade.

O treinamento foi estruturado em duas etapas, da seguinte forma:

- na primeira etapa, houve um estudo prévio do Manual do Recenseador; para o desenvolvimento dessa atividade, o treinando recebeu todos os instrumentos de coleta e foi orientado pelas instruções contidas no documento *Segunda prova piloto: roteiro para iniciação ao estudo do manual do recenseador* (1998), que estabeleciam um estudo dirigido em 17 etapas, onde, em cada uma, era solicitada a leitura de uma parte do Manual e a solução de um conjunto de exercícios; e

- a segunda etapa foi desenvolvida através de um treinamento presencial, realizado durante cinco dias, com carga horária total de 35 horas-aula; toda a atividade de ensino/aprendizagem teve o suporte do Manual do Instrutor, que apresentava a programação do treinamento, o roteiro das aulas e o gabarito das questões propostas no Caderno de Exercícios; o recurso instrucional básico foi a projeção de transparências; houve, também, a exibição de um vídeo sobre o Censo Demográfico; ao fim da atividade, foi feita uma avaliação da aprendizagem e solicitada uma avaliação de reação a cada um dos treinandos.

Todo esse programa foi antecedido por um treinamento interno em que, no período de 13 a 17 de julho de 1998, foram capacitados técnicos da DPE e da COI/DI, que participariam como instrutores e observadores da operação, além dos coordenadores do teste de distribuição prévia e autopreenchimento dos questionários, indicados pelas Unidades Regionais do Rio de Janeiro e São Paulo.

O treinamento para a equipe de campo realizou-se na semana de 27 de julho a 1º de agosto, com seis duplas de instrutores da DPE repassando as instruções em cada um dos municípios selecionados para a prova piloto, à exceção do Rio de Janeiro e São Paulo, onde não houve coleta de dados através de entrevistas.

Coleta dos dados

Como já observado, a coleta das informações da Segunda Prova Piloto foi realizada utilizando-se duas metodologias: a de preenchimento dos questionários pelo próprio informante e a da entrevista, através da visita do Recenseador aos domicílios.

Coleta por autopreenchimento

Todas as instruções necessárias à realização do teste de autopreenchimento estavam reunidas no folheto Orientações para a Realização do Teste de Autopreenchimento da Segunda Prova Piloto do Censo Demográfico 2000.

As áreas de coleta foram escolhidas, previamente, em condomínios, nos quais, através das informações prestadas por uma pessoa responsável – porteiro, síndico administrador etc. – foi realizada, por um funcionário do IBGE, a listagem de domicílios, utilizando-se a Folha de Coleta e as instruções do Manual do Recenseador, considerando cada área como um setor censitário. Nessa oportunidade foram também conhecidos o número de moradores de cada domicílio.

Para cada dez domicílios listados, nove responderam ao Questionário Básico e um ao Questionário da Amostra – fração de amostragem de 10%. Para tanto, foram entregues, à pessoa responsável os envelopes para encaminhamento aos domicílios, contendo cada um: carta ao morador, pesquisa de opinião, Manual de Preenchimento do Questionário Básico do Censo Demográfico 2000 ou Manual de Preenchimento do Questionário da Amostra do Censo Demográfico 2000 e o número de questionários suficientes para a coleta.

Os envelopes foram distribuídos no período de 24 a 28 de agosto de 1998, sendo estabelecido o dia 15 de novembro como o prazo máximo para que o funcionário do IBGE recolhesse, com a pessoa responsável, os questionários e a pesquisa de opinião preenchidos.

Em seguida, houve um trabalho de investigação do material preenchido para corrigir erros de seqüência ou falta de informação em quesitos de preenchimento obrigatório; para isso, o funcionário do IBGE devia fazer contato telefônico com os informantes.

Por fim, todos os erros encontrados, corrigidos ou não, foram relatados, para cada área selecionada para o teste, na Folha de Apuração de Erros por Setor, conforme orientações do Manual do Supervisor.

Coleta por entrevistas

A coleta da Segunda Prova Piloto teve início no dia 3 de agosto de 1998, estendendo-se até o final do mês. O trabalho de entrevista foi realizado por Recenseadores contratados para cada setor, enquanto a atividade de supervisão foi executada por técnicos lotados nas Unidades Regionais, de modo que cada Supervisor ficou responsável por, no máximo, cinco Recenseadores.

Os técnicos responsáveis pelo treinamento permaneceram nas Unidades Regionais, durante a primeira semana de coleta, para avaliar a funcionalidade, conteúdo e formato dos instrumentos de coleta. Durante a última semana da pesquisa, outros técnicos da sede do IBGE atuaram como observadores da prova piloto.

O Questionário Básico foi aplicado em 75% dos domicílios particulares e o Questionário da Amostra nos restantes 25%. A utilização dessa fração de amostragem foi a forma de tentar garantir suficiente quantidade de questionários e, conseqüentemente, de informações para análise.

A tabela, a seguir, apresenta algumas informações ligadas ao trabalho de coleta da Segunda Prova Piloto.

Tabela 5.1 - Técnicos envolvidos e setores investigados na Segunda Prova Piloto do Censo Demográfico, segundo as Unidades da Federação e municípios selecionados - 2000

Unidades da Federação e municípios selecionados	Setores		Recenseadores	Supervisores	Observadores
	Urbanos	Rurais			
Total	**75**	**15**	**84**	**18**	**37**
Paraíba	25	5	30	6	12
João Pessoa	15	0	15	3	6
Campina Grande	10	5	15	3	6
Minas Gerais	25	5	27	6	14
Belo Horizonte	20	0	17	4	9
Governador Valadares	5	5	10	2	5
Goiás	25	5	27	6	11
Goiânia	20	0	19	4	7
Rio Verde	5	5	8	2	4

Fonte: IBGE, Segunda Prova Piloto do Censo Demográfico 2000.

Diferentemente da prova piloto anterior, onde foram testados somente os questionários, a Segunda Prova Piloto fez uma simulação completa da coleta de dados. Assim, o trabalho contou, efetivamente, com os instrumentos já consagrados nesse tipo de operação, a saber: a Caderneta do Setor, a Folha de Coleta, a Folha de Domicílio Coletivo, e os dois tipos de questionários já descritos.

Em função disso, Recenseadores e Supervisores executaram todas as tarefas inerentes ao trabalho de coleta, desde a atividade de reconhecimento prévio do setor censitário e conferência de seu mapa, passando pela listagem de suas unidades e seleção dos domicílios da amostra, até a realização das entrevistas e do trabalho de supervisão.

Ao Recenseador coube, ainda, a elaboração de um relatório que, em sua primeira parte, fornecia informações sobre as entrevistas para avaliação de seu tempo de duração e, na segunda, preenchida ao fim do trabalho de campo, apresentava respostas a um conjunto de quesitos.

O Manual do Supervisor reunia todas as informações relativas ao trabalho a ser executado, com ênfase especial para a rotina de supervisão do trabalho do Recenseador que, em linhas gerais, passava pelas seguintes atividades:

- acompanhamento dos Recenseadores durante a primeira semana de coleta;

- exame, no escritório, do preenchimento dos instrumentos de coleta; e

- realização de reentrevistas.

Terminada a coleta, o Supervisor fazia a revisão dos lançamentos da Caderneta do Setor e o preenchimento dos seguintes documentos:

- Folha de Apuração Preliminar, cujos objetivos eram agilizar a divulgação dos resultados preliminares da população dos municípios e fornecer informações da coleta a serem avaliadas pela GTD;

- Folha Resumo do Setor dos Questionários Básico e da Amostra que congregava informações para formar a base de conferência dos totais de documentos remetidos; e

- Folha de Encaminhamento de Setores.

Em seguida, empastava todos os instrumentos de coleta e enviava-os à agência do IBGE. Devia remeter, também, o Relatório do Supervisor, que continha a Folha de Apuração de Erros por Setor e as respostas a 6 quesitos formulados em aberto.

5.2.4 Sistema de acompanhamento da coleta

A premissa básica do projeto SIGC para a Segunda Prova Piloto do Censo Demográfico 2000 era poder realizar o acompanhamento centralizado de toda a coleta e a obtenção dos resultados, durante e ao fim do desenvolvimento dos trabalhos de campo.

Além de permitir o acompanhamento centralizado de toda a coleta, a solução facilitou a análise dos dados coletados durante os trabalhos de campo e, ainda, a apresentação dos resultados de avaliação.

Os procedimentos adotados durante o desenvolvimento, a instalação e a utilização do Sistema de Acompanhamento da Coleta e Resultados Preliminares da Segunda Prova Piloto do Censo Demográfico 2000 serão descritos, a seguir.

5.2.4.1 Modelo conceitual

O objetivo era possibilitar o acompanhamento da coleta e a divulgação de resultados da Segunda Prova Piloto do Censo Demográfico 2000, com velocidade maior do que a apresentada nos censos anteriores. A solução encontrada pela Diretoria de Informática foi implantar um sistema nas agências envolvidas, com tecnologia de comunicação e transmissão das informações até as respectivas Divisões de Pesquisas, e daí para o equipamento central; tudo compatível com as necessidades apresentadas pelos técnicos da Diretoria de Pesquisas.

Processos envolvidos

Atualizar sistema – informar ao sistema dados referentes aos setores que estão entrando em coleta.

Replicar informações DIPEQ – replicar informações atualizadas localmente, para o servidor da DIPEQ.

Consultas agência – acompanhar, através do sistema, o andamento dos trabalhos da coleta no âmbito da agência.

Replicar informações ao servidor central – replicar informações para o servidor da DI.

Consultas DIPEQ – acompanhar, através do sistema, o andamento dos trabalhos da coleta.

Gerenciamento do sistema – gerenciar todo o processo de manutenção e utilização do sistema.

Consultas DI – gerar relatórios gerenciais e de controle.

Consultas arq-pesquisa – disponibilizar dados dos setores encerrados/liberados em formato “txt”.

5.2.4.2 Nível de desagregação e variáveis em estudo

O sistema foi dividido em dois módulos, de modo a facilitar sua utilização.

Módulo resultados preliminares

Este módulo possibilitou aos usuários do sistema a obtenção de informações relativas aos resultados preliminares da coleta, considerando as seguintes variáveis:

data de atualização do sistema;

domicílios particulares ocupados;

domicílios particulares fechados;

domicílios particulares de uso ocasional;

domicílios particulares vagos;

unidades de habitação em domicílios coletivos;

total de recusas;

total de homens e de mulheres – Questionário Básico;

total de homens e de mulheres – Questionário da Amostra;

E segundo os seguintes níveis de desagregação:

agência;

município ou conjunto de municípios;

distrito e subdistrito; e

setor.

Módulo acompanhamento da coleta

Este módulo permitiu acompanhar a coleta, segundo a localização dos setores, em área urbana ou rural, observando-se as seguintes informações:

total de setores em trabalho;

total de setores não iniciados;

total de setores interrompidos;

total de setores coletados; e

total de setores liberados.

5.2.4.3 Tecnologia envolvida

A tecnologia utilizada concentrou-se, basicamente, em solucionar duas questões do processo: a comunicação entre grupos e a comunicação com a rede de informática do IBGE.

Em relação à comunicação entre grupos, o sistema foi desenvolvido em ambiente Lotus Notes, o que permitiu a utilização de sua estrutura de segurança, que possibilita atribuir restrições a cada usuário cadastrado e requerer a sua identificação antes de utilizar o sistema.

Na comunicação efetiva com a rede, foram aproveitadas a estrutura institucional existente e vários equipamentos disponíveis, possibilitando reduzir o custo da solução. Como a rede institucional já atingia as DIPEQs, que tinham o Lotus Notes instalado, a inovação, em termos de *hardware*, consistiu em conectar, remotamente, as agências à rede existente, através de linha discada, via acesso *Dial-Up*. Assim, as agências tornaram-se nós da rede institucional, permitindo a utilização do presente sistema.

Configuração de *hardware* utilizada

A configuração de *hardware* utilizada estava dividida em dois grupos, a saber: o instalado nas agências; e o preexistente nas DIPEQs, acrescido dos requisitos necessários para viabilizar a conexão com as agências.

Configuração nas agências

O Quadro 5.1 mostra as características dos equipamentos instalados nas agências envolvidas nos trabalhos de coleta da Segunda Prova Piloto.

Quadro 5.1 - Características dos equipamentos instalados nas agências envolvidas na Segunda Prova Piloto do Censo Demográfico - 2000

Unidades da Federação e agências	Equipamento				Linha telefônica	
	Processador	Disco	Memória	Fax modem		
Minas Gerais						
Noroeste	Pentium 166	3.2 Gb	32 Mb	33 600 Bps	Analógica 24 576 Bps	Conec/aprox.
Acaiaca	Pentium 100	1.7 Gb	32 Mb	Rede local	Circ. dedicado	
Governador Valadares	Pentium 166	3.2 Gb	32 Mb	33 600 Bps	Digital 24 576 Bps	Conec/aprox.
Goiás						
Goiânia	Pentium 166	3.2 Gb	32 Mb	57 344 Bps	Analógica 38 400 Bps	Conec/aprox.
Rio Verde	Pentium 166	3.2 Gb	32 Mb	57 344 Bps	Analógica 19 200 Bps	Conec/aprox.
Paraíba						
João Pessoa	Pentium 166	3.2 Gb	32 Mb	Rede local	Circ. dedicado	
Campina Grande	Pentium 166	3.2 Gb	32 Mb	33 600 Bps	Analógica 38 400 Bps	Conec/aprox.

Configuração dos equipamentos nas DIPEQs

Além dos microcomputadores já em uso, foi instalado em cada DIPEQ um *modem* Motorola de 56 Kbpsm, conectado a uma linha telefônica e a uma porta no servidor de terminais já existente.

Configuração de software utilizada

Nas agências, foi utilizado o sistema operacional Windows 95, o correio eletrônico Lotus Notes e, ainda, o protocolo para integração com a rede.

Lotus Notes

O Lotus Notes é um *software* para comunicação de grupos baseado em redes, que permite às pessoas dentro de uma organização, independentemente de sua localização geográfica, criar, organizar, ter acesso e compartilhar diferentes tipos de informações. Utilizando-se do conceito de réplicas, o Lotus Notes cuida da transferência de informações entre as unidades envolvidas no processo, garantindo totalmente a integridade das informações.

Com arquitetura cliente-servidor, e suportando intercâmbio dinâmico de dados e objetos vinculados e incorporados, permitiu aos usuários integrar uma variedade de aplicativos Windows. Sua escolha deveu-se ao fato de ser o Lotus Notes a tecnologia que mais se adaptava, na época, às finalidades do projeto.

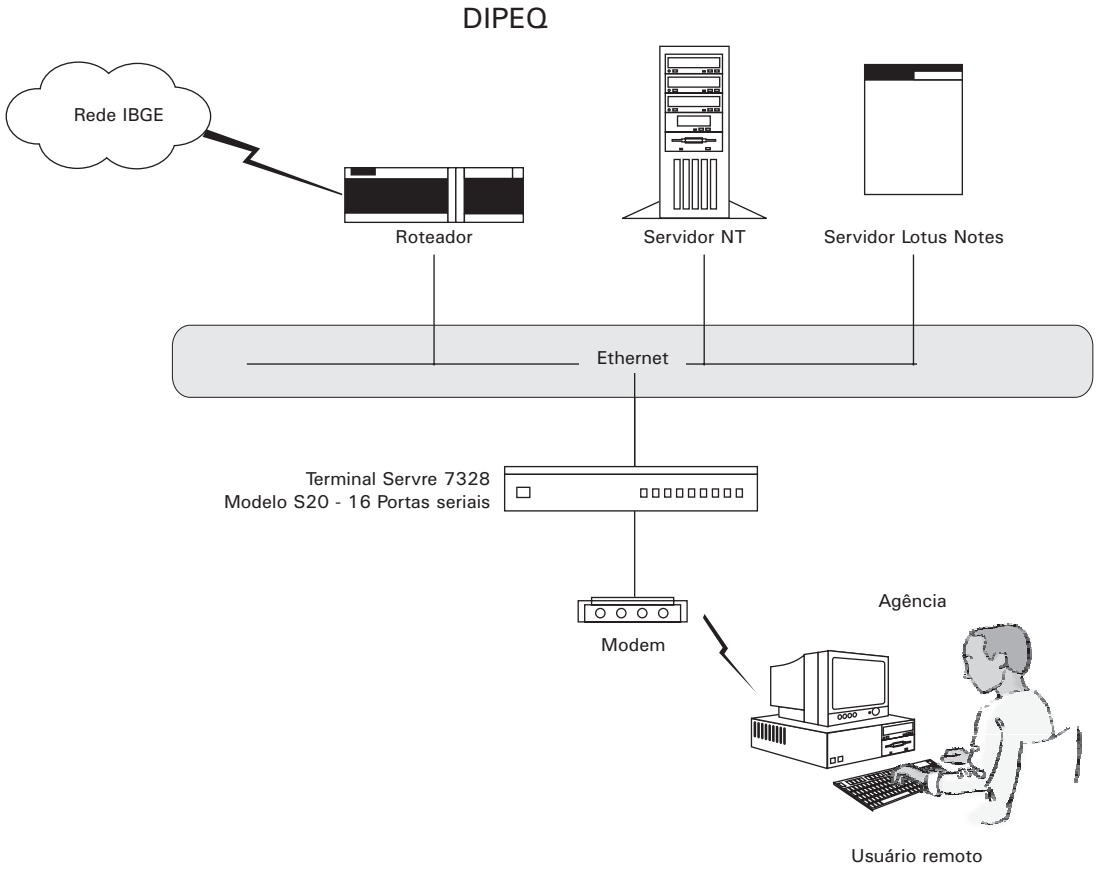
Integração com a rede IBGE

A integração das agências com a rede do IBGE foi feita através da conexão *Dial-up*. Essa conexão permitia ao usuário conectar-se a um computador remoto, por meio de um *modem*, através de uma linha telefônica pelo protocolo utilizado.

Quando conectado, o computador remoto passa a atuar como um nó da rede local ao adquirir um endereço para a rede local, com todos os privilégios que foram atribuídos para esse usuário.

A agência executava a entrada de dados localmente, e através da conexão da rede, transmitia os dados para o servidor localizado na DIPEQ. A figura 5.1 ilustra a referida conexão.

Figura 5.1 – Esquema de integração das agências à rede IBGE



5.2.4.4 Instalação e utilização do sistema

Conforme premissa básica do sistema, o mesmo foi instalado nas agências envolvidas, seguindo o modelo básico de máquina local, com ligação via *modem*. As modificações ocorreram nas agências Leste, Centro Sul e Oeste e na agência João Pessoa.

Devido ao fato de que as agências Leste, Centro Sul e Oeste encontravam-se no mesmo prédio e no mesmo andar, optou-se por configurar um único micro para envio das informações referentes às três agências. Como o local possuía uma infra-estrutura de rede, que permitia uma ligação direta ao prédio da DIPEQ/MG, através de um circuito de 19 200 bps, não foi utilizada a conexão *Dial-up*. Avaliou-se ser este um bom teste para o sistema devido a sua situação peculiar.

Na agência João Pessoa, também não foi utilizada a conexão *Dial-up*, pois possuía uma infra-estrutura de rede conectada à DIPEQ/PB.

Quanto à utilização do sistema, o mesmo não apresentou qualquer dificuldade de uso, ainda que alguns usuários tenham declarado não ter nenhuma prática no tocante ao uso de microcomputadores. Após receberem as instruções de uso e através de demonstrações práticas, os usuários se mostraram aptos a estabelecer a conexão *Dial-up* e a utilizar o sistema sem qualquer problema.

5.2.5 Avaliação da operação

A seguir, apresenta-se um breve resumo dos principais pontos referentes às avaliações do trabalho de coleta da Segunda Prova Piloto e do sistema informatizado utilizado para seu acompanhamento.

Trabalho de coleta

A tabela seguinte apresenta o tempo médio e o total de entrevistas, por tipo de questionário e número de moradores, para as Unidades Estaduais da Paraíba, Minas Gerais e Goiás.

Tabela 5.2 - Duração média e número de entrevistas e número de moradores, por Unidades da Federação, segundo o tipo de questionário utilizado na Segunda Prova Piloto do Censo Demográfico - 2000

Tipo de questionário	Número de moradores	Paraíba		Minas Gerais		Goiás	
		Tempo médio (min)	Total de entrevistas	Tempo médio (min)	Total de entrevistas	Tempo médio (min)	Total de entrevistas
Básico	1	7	275	9	322	7	373
	2	9	577	9	594	9	659
	3	10	887	10	683	9	827
	4	12	1 090	12	704	10	913
	5	13	764	13	495	11	494
	6	14	355	16	235	12	172
	7	17	176	14	106	15	75
	8	17	86	16	55	16	22
	9 ou mais	26	107	23	49	18	27
Amostra	1	17	82	25	107	16	121
	2	25	206	30	203	24	224
	3	30	296	36	207	25	261
	4	33	310	37	240	28	305
	5	38	274	42	135	29	162
	6	49	102	44	66	36	58
	7	51	61	49	31	40	21
	8	47	40	46	15	44	9
	9 ou mais	70	36	58	13	56	8

Fonte: IBGE, Segunda Prova Piloto do Censo Demográfico 2000.

Seguem alguns pontos observados no trabalho de avaliação:

- o Questionário da Amostra, mesmo simplificado e com seu desenho melhorado, foi considerado inadequado ao autopreenchimento, face à complexidade de comandos e conceitos envolvidos. Também o Questionário Básico apresentou problemas para o autopreenchimento, especialmente em alguns quesitos do bloco Características do Domicílio e nas perguntas de educação para a pessoa responsável pelo domicílio;

- quanto aos quesitos de características do domicílio, ficou nítido que vários blocos de quesitos com os comandos exigindo troca de colunas não era aconselhável em termos de formatação;

- a pergunta de cobertura para eventual omissão de moradores estava mal formulada e acarretou muitos erros, especialmente para o autopreenchimento, porque seu texto estava dirigido diretamente ao Recenseador, o que não era a situação dos informantes;

- no bloco de características gerais, ficou nítido que os informantes do autopreenchimento não compreenderam o enunciado e objetivo do quesito sobre "origem", deixando muitos deles em branco; houve também nítida dificuldade de compreensão dos informantes, quanto à numeração das famílias;

- no bloco de educação dois problemas se destacaram: a pergunta do grau para quem freqüenta e já freqüentou escola apresentada em um único quesito gerou confusão; a outra dificuldade ocorreu na conversão para o sistema seriado, de cursos freqüentados em modalidade não seriada;

- a dificuldade do bloco de nupcialidade concentrou-se no quesito sobre a natureza da última união; muitas das pessoas que já viveram em união mas, na data de referência, não se encontravam nesta situação, não tiveram a situação anterior informada;

- o problema no bloco de fecundidade foi comum aos dois métodos de coleta; a pergunta sobre migração internacional foi respondida erradamente para muitas pessoas, por se encontrar absolutamente fora do contexto do bloco de investigação; e

- no bloco de mão-de-obra, percebeu-se que o emaranhado de perguntas e comandos para captar e separar, em poucos quesitos, as pessoas ocupadas na semana e ano de referência, gerou erros nos dois métodos de entrevista com predominância do autopreenchimento; ficou claro também que a pergunta sobre o número de pessoas trabalhando no empreendimento não devia ter sido aplicada indistintamente a todas as pessoas ocupadas, sendo aconselhável que se restringisse aos empregadores.

Os relatórios de avaliação da segunda prova piloto do Censo Demográfico 2000 indicados na lista de Referências ao final do capítulo, apresentam as demais conclusões referentes à avaliação do trabalho de coleta de dados.

Sistema de acompanhamento da coleta

Os analistas responsáveis pelo desenvolvimento do sistema concluíram que os recursos utilizados durante a Segunda Prova Piloto não suportariam a grande operação do Censo Demográfico 2000. Ocorreram vários problemas na operação de réplicas, inviabilizando a obtenção de informações sobre o acompanhamento da coleta, em tempo adequado, consolidadas para as Unidades Regionais e o Brasil.

Ficou claro, também, que havia uma demanda muito grande por informações gerenciais em níveis de agregação diversos – municípios, agências etc. – com a geração dos dados em tabelas e gráficos de fácil observação.

Foi detectada a necessidade de separação das informações de acompanhamento da coleta e de resultados para a análise da situação da coleta relativamente a resultados de pesquisas passadas, como o Censo Demográfico 1991 e a Contagem da População 1996.

Uma outra observação importante foi a necessidade da introdução de modificações no sistema, tendo em vista a implementação mais adequada de críticas preliminares de entrada de dados.

5.3 Prova Piloto conjunta dos países do Mercosul ampliado[1]

A Primeira Prova Piloto Conjunta dos Países do Mercosul Ampliado foi realizada nos meses de outubro e novembro de 1998, em áreas de fronteira. A prova aconteceu simultaneamente em duas áreas: entre o Brasil (Foz do Iguaçu) e a Argentina (Puerto Iguazú) e entre o Brasil (Corumbá) e a Bolívia (Puerto Quijarro). A participação da equipe de técnicos brasileiros contou com o apoio financeiro do Fundo de População das Nações Unidas – FNUAP, que custeou a prova, sendo esta uma das atividades do Projeto Incorporação de Novas Metodologias e Integração com os Países do Mercosul no Censo Demográfico 2000, projeto conjunto entre os países do Mercosul, Bolívia e Chile.

Esta primeira prova tinha como objetivos testar:

- os conteúdos comuns dos questionários de cada país;

- as técnicas usadas no treinamento e capacitação;

- as metodologias de processamento e codificação assistida com relação às nomenclaturas utilizadas;

- o desenho da base de dados e a produção do plano tabular comum; e

- a introdução de novas metodologias.

O cumprimento das duas primeiras metas foi avaliado mediante uma metodologia de observação não-participativa, com o monitoramento por equipes de campo dos países dela participantes e representantes de outros países do Mercosul. Esta tarefa contribuiu para enriquecer a experiência, e se constituiu em um dos eixos principais para o planejamento conjunto entre os órgãos de estatística dos países.

De acordo com o que foi estabelecido, a observação não-participativa no campo foi direcionada para a avaliação de aspectos ligados à implementação dos conceitos e procedimentos, abrangendo também a observação do desempenho dos Recenseadores.

Quatro observadores brasileiros participaram da prova – dois especialistas do IBGE, que viajaram desde o Rio de Janeiro até Porto Iguaçu e Corumbá, e os outros dois vieram das equipes de coordenação e capacitação das sedes de Curitiba e do Mato Grosso do Sul. A participação estrangeira esteve representada pela Argentina, com dois observadores vinculados à coordenação do Censo do

[1] Os países membros do Projeto Censo Comum no Mercosul Ampliado são Argentina, Brasil, Paraguai, Uruguai, Bolívia e Chile.

Instituto Nacional de Estadísticas y Censos – INDEC, em Foz do Iguaçu, e pela Bolívia, com dois observadores do Instituto Nacional de Estadística de Bolívia – INE, em Corumbá. A agência do IBGE, em Foz do Iguaçu, serviu de base de operações para a equipe brasileira. Os contatos entre as instituições foram realizados pelos gerentes das equipes, que também foram responsáveis pelos contatos com os municípios e a imprensa local. Para divulgar a prova, foram organizadas duas reuniões com a imprensa: uma em Puerto Iguazú e outra em Corumbá, ocasião em que foi marcado oficialmente o começo da prova piloto.

Os resultados da Primeira Prova Piloto Conjunta foram de extrema relevância para o processo de planejamento e realização dos censos, em cada país participante. A experiência mostrou que, freqüentemente, os mesmos conceitos, definições e classificações são interpretados, no campo, de maneira diferenciada em cada país. Assim sendo, foi solicitado que os relatórios dos observadores citassem as situações em que isso acontecia. Os relatórios não abordaram somente conceitos e definições que, na prática, foram operacionalizados diferentemente em cada país, mas também problemas de organização do questionário.

Os observadores brasileiros fizeram as seguintes observações acerca da aplicação do questionário do censo argentino:

- a coleta dos dados foi feita pelos professores da própria localidade, o que facilitou muito a receptividade da população;

- segundo o Manual do Recenseador, os Recenseadores não poderiam renunciar à tarefa para a qual eram designados, já que era dever cívico;

- um aspecto positivo do questionário foi o preenchimento da lista de pessoas para conhecer a estrutura da família, antes de se investigar as características da "VIVIENDA – hogar";

- na parte referente a características das pessoas só era preciso preencher com o número de ordem, que a pessoa recebeu na lista de moradores;

- foi sugerido que o desenho do questionário argentino fosse alterado, de modo a verticalizar, ao máximo, os blocos de perguntas, evitando assim a omissão de quesitos; e

- a opção de ignorado, em alguns quesitos, poderia ser eliminada para evitar que houvesse, conforme foi observado no campo, uma concentração de respostas nessa opção.

Ainda nessa etapa, os observadores argentinos reportaram sobre o censo brasileiro os seguintes aspectos: a população foi muito receptiva em relação ao trabalho dos técnicos do IBGE; o formulário avançou substancialmente, considerando seu formato e seu desenho, em comparação com o de 1991; as seqüências, marcadas com setas, facilitaram a tarefa do Recenseador; os esclarecimentos incorporados ao questionário ajudaram tanto ao Recenseador quanto ao entrevistado; e a lista de moradores do domicílio foi obtida facilmente a partir da pergunta "quem vive neste domicílio", realizada pelo Recenseador.

O aspecto negativo ficou por conta do questionário muito extenso, sendo por isso sugerido que alguns blocos fossem reduzidos, como por exemplo os de bens do domicílio e de trabalho, bem como a lista de relação de parentesco com a pessoa responsável.

Na área de fronteira entre o Brasil e a Bolívia, foram observadas diferenças no processo de capacitação da equipe de censitários. O Brasil apresentou uma boa metodologia de treinamento antecipado para a prova, que incluiu uma etapa de avaliação prévia do grupo que atuou em campo. A Bolívia empregou a metodologia de capacitação utilizada no Censo de 1992, com a realização de exposições e material de apoio. Além disso, os técnicos brasileiros fizeram algumas observações sobre o censo boliviano, a saber:

- a coleta de dados foi realizada por estudantes secundaristas, pelo exército e por professores;

- o Recenseador percorreu sua área de atuação previamente à coleta, acompanhado do supervisor, a fim de verificar a correspondência da área com o mapa;

- segundo o manual, o Recenseador não poderia renunciar às suas funções, salvo por alguma causa plenamente justificada e aprovada pelo Instituto Nacional de Estadística de Bolívia – INE: o trabalho era dever cívico;

- foram feitas algumas sugestões para que todo quesito tivesse um sistema de comando para facilitar a entrevista, e também que os saltos entre os blocos fossem destacados;

- o Manual do Recenseador não esclarecia quais cursos deveriam ser considerados como freqüência à escola;

- as perguntas, de modo geral, foram lidas como figuram no questionário; e

- os técnicos brasileiros receberam todo o apoio necessário à realização de seu trabalho em campo.

Cabe ainda ressaltar que todos os países participantes escolheram a metodologia de coleta baseada na entrevista domiciliar, que foi feita por Recenseadores treinados para tal fim.

Em ambas as áreas foi observada uma grande semelhança entre as características das pessoas moradoras nos dois lados da fronteira. Esta semelhança era maior que aquela observada com pessoas residentes em outras localidades do próprio país de origem. Seus modos de viver eram parecidos, tanto nas relações familiares quanto no tipo de trabalho. Apesar do fato da população ser muito parecida e dos questionários serem similares, ainda foram observadas diferenças devido à linguagem, que fez com que as interpretações de um lado e de outro da fronteira fossem distintas.

A partir da análise das divergências relatadas, foi sugerido que se incorporassem especificações mais detalhadas nas definições aos manuais de instrução. Ao longo de todo o processo, foi possível constatar que a Prova Piloto Conjunta constituiu uma experiência rica e produtiva, permitindo o aprimoramento da abordagem e da forma de se investigar nos questionários, e, principalmente, representando um evento de significativo intercâmbio entre os países integrantes do projeto.

5.4 Censo Experimental

O Censo Experimental tem com o objetivo simular a operação censitária, através de um protótipo da pesquisa, implementado em áreas selecionadas, de modo a permitir a avaliação de todo o processo de trabalho, a saber:

- seleção de pessoal e treinamento das equipes;

- coleta das informações, incluindo material acessório, como lápis, borracha e prancheta;

- sistemas de acompanhamento da coleta e de pagamento de Recenseadores;

- digitalização dos questionários e reconhecimento ótico de caracteres; e

- crítica, imputação e tabulação dos dados.

Assim, a realização do Censo Experimental, por levar a campo, efetivamente, a última versão de todos os procedimentos que seriam adotados no Censo 2000, inclusive alguns ainda não testados nas duas provas piloto, ensejou a última possibilidade de ajustes, com vistas ao aprimoramento da pesquisa futura.

5.4.1 Âmbito da pesquisa

A orientação que norteou o processo de escolha dos municípios, com vistas à realização do Censo Experimental 1999, foi a mesma dos censos experimentais anteriores, a saber: a proximidade da sede do IBGE, no Rio de Janeiro; a existência de meios de comunicação adequados para a operação; e a presença de atividade econômica diversificada. Esses fatores têm feito com que as escolhas venham recaindo em municípios de porte médio do Estado de São Paulo, conforme ocorreu, em 1967, com Presidente Prudente, em 1979, com Taubaté e, em 1988, com o Município de Limeira.

Houve, ainda, a preocupação de tentar, ao máximo, reproduzir nessa operação experimental, as dificuldades inerentes ao processo definitivo da pesquisa, como condições climáticas adversas, precária infra-estrutura de transportes e comunicações etc. Assim, optou-se também pela inclusão de um município da Região Norte que, de alguma forma, possibilitasse a avaliação dessas dificuldades durante a coleta. Desse modo, os municípios selecionados foram Marília (SP) e Bonito (PA).

O Censo Experimental foi realizado nos meses de agosto e setembro de 1999, mediante a aplicação de dois modelos de questionários – o Questionário Básico e o Questionário da Amostra. O Questionário da Amostra foi aplicado em 10% dos domicílios particulares no Município de Marília, uma vez que possuía população maior que 15 000 habitantes. Em Bonito, cuja população encontrava-se abaixo desse número, a fração de amostragem utilizada foi de 20%. Nos dois casos, o Questionário Básico foi aplicado aos demais domicílios não selecionados para a amostra.

A Base Operacional Geográfica – mapas e descrições dos setores – utilizada para o trabalho de campo, nos dois municípios, foi a preparada para o Censo 2000, já incorporadas todas as atualizações. Foram definidos 300 setores para o Município de Marília, sendo 275 na área urbana e cinco na área rural, e 13 setores para o Município de Bonito, sendo dois na área urbana.

5.4.2 Características investigadas

Os dois tipos de questionário investigaram características do domicílio e do morador.

O Questionário Básico apresentou o seguinte conteúdo:

- características do domicílio – espécie, tipo, condição de ocupação do domicílio e do terreno onde se localizava, forma de abastecimento e canalização de água, número de banheiros, existência de sanitário, tipo de ligação do escoadouro e destino do lixo; e

- características do morador – para todos os moradores: sexo, relação com a pessoa responsável pelo domicílio, idade e se sabe ler e escrever; somente para o responsável pelo domicílio ou individual em domicílio coletivo: curso mais elevado que freqüentou com conclusão de, pelo menos, uma série, última série concluída com aprovação, e rendimento bruto, auferido em julho de 1999.

O Questionário da Amostra investigou em:

- características do domicílio – além do conteúdo investigado pelo Questionário Básico, foram, ainda, incluídos: número de cômodos, número de cômodos servindo de dormitório, existência de iluminação elétrica, rádio, geladeira ou *freezer*, videocassete, máquina de lavar roupa, forno de microondas, linha telefônica instalada, microcomputador e quantidade de televisores, automóveis para uso particular, aparelhos de ar condicionado; e

- características do morador – além do que foi pesquisado pelo Questionário Básico, foram adicionados os seguintes pontos: relação com a pessoa responsável pela família, número da família, cor ou raça, religião, deficiência física ou mental, migração, escolaridade, nupcialidade, trabalho e rendimento, e fecundidade.

Foi pesquisada, também, através da Folha de Coleta e da Caderneta do Setor, a caracterização de cada logradouro ou trecho de logradouro percorrido no setor, correspondendo à "existência ou não de identificação", "iluminação pública" e calçamento/pavimentação".

5.4.3 Treinamento

O treinamento do Censo Experimental foi desenvolvido em diversas datas – ver Quadros 5.2 e 5.3 – no período de 17 de junho a 30 de julho de 1999, com o objetivo de capacitar os treinandos para aplicarem os conceitos e procedimentos estabelecidos, bem como habilitar, como instrutores, aqueles que iriam repassar o treinamento nas etapas subseqüentes.

A implementação do treinamento baseou-se nas seguintes estratégias:

- capacitação didático-pedagógica para os instrutores do IBGE;

- treinamento para os Supervisores sobre a Base Operacional Geográfica e sobre os sistemas administrativos;

- realização de atividades de auto-instrução;

- realização de cursos presenciais de curta duração, com ênfase na participação ativa dos treinandos e voltados, essencialmente, para a operacionalização dos conceitos do Censo;

- utilização do recurso de vídeo-aula;

- utilização de manuais técnicos com tratamento pedagógico e outros materiais impressos; e

- aplicação de um sistema de acompanhamento e avaliação do treinamento.

Inicialmente, houve a preparação didático-pedagógica dos técnicos da DPE e das Unidades Regionais do Pará e de São Paulo, que participaram como instrutores da 1ª e da 2ª etapas do treinamento, realizadas através do Curso de Capacitação de Instrutores para o Censo 2000, ministrado, durante três dias, pela ENCE.

Dando seguimento ao processo de capacitação das equipes, foi estabelecido um cronograma para o treinamento do Censo Experimental, obedecendo às seguintes etapas:

1ª etapa – treinamento administrativo, técnico-operacional e sobre a Base Operacional Geográfica, para os Coordenadores Técnicos e de Subárea de Marília e Bonito e para os ACMs de Marília, realizado nas instalações das URs; participaram, também, desta etapa, outros técnicos das URs de São Paulo e Pará;

2ª etapa – treinamento para os Supervisores de Marília e Bonito e para o ACM de Bonito:

• 1ª fase – Base Operacional Geográfica e Sistema Administrativo para a contratação, dispensa e pagamento dos Recenseadores;

• 2ª fase – técnico-operacional; e

• 3ª etapa – treinamento técnico-operacional para os Recenseadores.

As alternativas de carga horária estabelecidas para o treinamento foram as seguintes:

- nas duas primeiras etapas, 32 horas-aula;

- na terceira etapa, 20 horas-aula, com o auxílio de vídeo-aula, no Município de Bonito e em seis turmas de Marília;

- 10 horas-aula, com o auxílio de vídeo-aula, em uma turma de Marília; e

- 24 horas-aula, sem o recurso de vídeo-aula, em Marília, para duas turmas de treinandos.

O treinamento técnico-operacional foi desenvolvido, tendo como base os seguintes recursos didáticos, apresentados no anexo de CD-ROM:

- TR 01 – Manual do Instrutor – apresentava todas as orientações necessárias ao trabalho do instrutor, oferecendo o passo a passo para sua tarefa;

- TR 02 – Roteiro de Estudo para o Recenseador – tinha o objetivo de orientar o treinando, durante a auto-instrução, no estudo do Manual do Recenseador;

- TR 03 – Teste Inicial – tinha por finalidade habituar os treinandos com os conceitos e procedimentos de coleta definidos para o Censo;

- TR 04 – Caderno de Exercícios – consistia num conjunto de exercícios, tendo em vista fixar os conceitos, as formas de preenchimento dos instrumentos e os procedimentos de campo estabelecidos para a coleta do censo;

- TR 06 – Teste Final A e TR 07 – Teste Final B – respectivamente, faziam a verificação do aprendizado dos treinandos para Supervisores e de Recenseadores;

- TR 08 – Gabarito do Teste Final A e TR 09 – Gabarito do Teste Final B – fornecia os critérios para correção dos testes, bem como as formas de avaliação dos treinandos;

- TR 1.10 – Avaliação do Treinamento – Treinandos – foi preenchida por uma amostra de treinandos das diversas fases do treinamento;

- TR 1.11 – Avaliação do Treinamento – Instrutor – feita por todos os instrutores; e

- Vídeo-aula, álbum seriado, instrumentos de coleta – CD 1.01, CD 1.02, CD 1.03, CD 1.07 e CD 1.09 – e instrumentos de coleta ampliados do CD 1.03 e CD 1.07.

As técnicas de ensino-aprendizagem, implementadas durante o treinamento técnico-operacional, foram as seguintes: exposição oral, leitura dirigida, discussão em duplas, discussão em pequenos grupos e dramatização.

Na 2ª e na 3ª etapas, as atividades de instrução foram desenvolvidas em escolas cedidas ao IBGE, sendo os instrutores da última etapa, os técnicos contratados que, mais tarde, atuaram como Supervisores no Censo Experimental.

Todas as etapas do treinamento foram acompanhadas por técnicos da COC, da DPE e das Unidades Regionais Pará, São Paulo, Amazonas, Bahia, Rio Grande do Sul e Distrito Federal que, como observadores, avaliaram os recursos didáticos – auto-instrução, manuais do instrutor, recenseador e do supervisor, vídeo-aula, álbum seriado e instrumentos ampliados – o desempenho técnico dos instrutores, e a sistemática do treinamento – carga horária, seqüência das aulas etc.; esses mesmos pontos foram objeto de avaliação por parte dos treinandos. A assimilação do conteúdo do treinamento foi verificada através da aplicação do Teste Final, que fez a aferição objetiva do conhecimento dos treinandos.

Os Quadros 5.2 e 5.3, a seguir apresentados, permitem uma melhor visualização do desenvolvimento do treinamento do Censo Experimental.

Quadro 5.2 - Esquema de treinamento do Censo Experimental no Município de Marília - São Paulo - 1999

Treinandos		Instrutores		Treinamento	Período	Locais	Observadores	
Grupo	Número	Área	Número				Area	Número
Técnicos da unidade regional	15	ENCE	1	Didático-pedagógico	17/06 a 19/06	São Paulo		
Coordenador técnico, Coordenador de área, subárea, ACMs e técnicos da unidade regional	15	DPE	2	Técnico operacional, administrativo, base operacional	29/06 a 02/07	São Paulo	COC/UR	1
Supervisores	27	UR-ACMs	3	1ª fase administrativa/ base operacional	06/07 e 07/07	Marília	COC/UR/SP	2
				Auto-instrução	08/07 a 12/07			
		UR-ACMs	3	2ª fase técnico-operacional	13/07 a 16/07	Marília	COC/DPE	2
				Auto-instrução	12/07 a 25/07			
Recenseadores	214	UR-supervisores	18	Técnico-operacional	26/07 a 30/07	Marília	COC/DPE/Urs São Paulo, Distrito Federal, Bahia e Rio Grande do Sul	9

Nota: No treinamento dos Recenseadores, foram constituídas nove turmas.

Quadro 5.3 - Esquema de treinamento do Censo Experimental no Município de Bonito - Pará - 1999

Treinandos		Instrutores		Treinamento	Período	Locais	Observadores	
Grupo	Número	Área	Número				Área	Número
Técnicos da unidade regional	15	ENCE	1	Didático-pedagógico	17/06 a 19/06	Belém		
				Auto-instrução	21/06 a 28/06			
Coordenador técnico, Coordenador de área, subárea, e outros técnicos	5	DPE	1	Técnico operacional, administrativo, base operacional	29/06 a 02/07	Belém	COC	1
ACM/Supervisores	3	UR	2	1ª fase administrativa/ base operacional	06/07 e 07/07	Bonito		
		UR	2	2ª fase técnico-operacional	13/07 a 16/07	Bonito	DPE	1
				Auto-instrução	12/07 a 25/07			
Recenseadores	11	ACM/supervisores	2	Técnico-operacional	26/07 a 30/07	Bonito	DPE/Urs Pará e Amapá	3

Nota: No treinamento dos Recenseadores, foram constituídas nove turmas.

5.4.4 Coleta dos dados

A coleta do Censo Experimental começou, nos dois municípios, no início de agosto de 1999, um ano antes da data prevista para o início da coleta do Censo Demográfico 2000. O processo de coleta utilizado foi o de entrevista presencial, tendo sido, também, testada a investigação através do autopreenchimento. Com duração prevista para dois meses, essa etapa da operação foi concluída no prazo estipulado nos dois municípios escolhidos.

Foram contratados, através de teste seletivo, 208 Recenseadores e 27 Supervisores para o Município de Marília. Por meio desta mesma modalidade, foram contratados 11 Recenseadores, dois Supervisores e um Agente Censitário Municipal para o Município de Bonito. O ACM de Marília era do quadro do IBGE, os Supervisores foram contratados como mensalistas, e os Recenseadores receberam pagamento por tarefa/produção.

5.4.4.1 Instrumentos de coleta

A seguir, apresentam-se os instrumentos utilizados no Censo Experimental durante a coleta dos dados. Para informações mais detalhadas, consultar fontes de informação na lista de Referências ao final do capítulo ou apresentadas em anexo.

CD 1.01 – Questionário Básico

Utilizado para o registro das características do domicílio e dos seus moradores, na data de referência, em cada unidade residencial ocupada, não selecionada para a amostra.

CD 1.02 – Questionário da Amostra

Utilizado para o registro das características do domicílio e dos seus moradores, na data de referência, em cada unidade residencial ocupada, selecionada para a amostra.

CD 1.03 – Folha de Domicílio Coletivo

Utilizada para listar as famílias e pessoas sós recenseadas em cada domicílio coletivo; serviu, ainda, para a seleção das unidades nas quais foi aplicado o Questionário da Amostra.

CD 1.06 – Caderneta do Setor

Apresentava o mapa e a descrição do setor e destinava-se ao registro do resumo das informações coletadas no setor e à guarda das Folhas de Coleta e das Folhas de Domicílio Coletivo.

CD 1.07 – Folha de Coleta

Utilizada para o registro das unidades residenciais e não-residenciais existentes no setor e do número de moradores em cada domicílio ocupado; serviu também para a caracterização dos logradouros e seleção dos domicílios particulares, nos quais seria aplicado o Questionário da Amostra.

CD 1.09 – Manual do Recenseador

Fornecia o suporte para o efetivo trabalho de coleta. Nele, estavam os conceitos, as definições e as normas de preenchimento dos instrumentos de coleta, bem como as demais orientações necessárias ao desempenho das tarefas do Recenseador.

CD 1.10 – Manual do Supervisor

Apresentava as instruções necessárias ao trabalho do Supervisor, começando pelo conhecimento das áreas de trabalho, passando pela tarefa de revisão da Base Operacional Geográfica e da rotina de crítica do trabalho do Recenseador, até o encerramento da coleta, preparo e remessa do material.

Agenda do Supervisor

Seu preenchimento objetivava montar um histórico do trabalho de cada Recenseador, organizando suas visitas ao posto de coleta para entrega do material e mantendo um registro das ocorrências identificadas durante a aplicação da rotina de supervisão. Continha, ainda, as informações semanais para lançamento no Sistema de Informações Gerenciais da Coleta.

CD 1.11 – Manual do Agente Censitário Municipal

Resumia as tarefas do ACM, ressaltando sua função de gerente do posto de coleta, com a descrição de todas as atividades técnicas e administrativas sob sua responsabilidade.

CD 1.12 – Manual de Coordenação

Esclarecia as funções técnicas e administrativas referentes ao trabalho a ser desempenhado pelo Coordenador Técnico e pelo Coordenador de Subárea, apresentando, detalhadamente, cada uma das tarefas que lhes foram atribuídas para a condução dos trabalhos censitários na Unidade Regional.

5.4.4.2 Teste de autopreenchimento

O teste de autopreenchimento, realizado em cerca de 10 000 domicílios, em cada um dos municípios do Rio de Janeiro e de São Paulo, como parte da Segunda Prova Piloto, apontou a inadequação dessa forma de investigação, sobretudo para o Questionário da Amostra.

Para o Questionário Básico, optou-se por mais um teste no Censo Experimental, considerando alguns ajustes e restrições em relação ao método adotado para a Segunda Prova Piloto.

No Censo Experimental, o autopreenchimento foi direcionado às regiões ou áreas onde o nível médio de instrução dos responsáveis pelos domicílios estivesse em torno do 2º grau completo ou superior, aplicável apenas nos domicílios onde o Recenseador encontrasse dificuldades para realizar a entrevista da forma convencional.

Desse modo, o questionário com as instruções e o envelope – para proteger o sigilo das informações prestadas – eram deixados no domicílio e recolhidos em uma segunda passagem do Recenseador. Para isso, foram selecionados 27 setores do Município de Marília, com base na média de anos de estudo dos responsáveis pelos domicílios, informada na Contagem da População 1996.

5.4.4.3 O trabalho de campo

São apresentadas, a seguir, de maneira geral, as atribuições dos membros da equipe de campo durante a realização do Censo Experimental.

Atribuições do Recenseador

Ao Recenseador, antes de dar início à coleta dos dados, coube percorrer o setor onde iria trabalhar, de modo a conhecer, sobretudo, os seus limites. Nessa oportunidade, verificava as diferenças entre o mapa do setor e o que foi observado no percurso prévio, comunicando-as ao seu Supervisor.

Durante o trabalho de coleta, o Recenseador realizava a cobertura do setor, fazendo a listagem de suas unidades e as entrevistas domiciliares, com base nas orientações do CD 1.09. A listagem consistia no cadastramento, no CD 1.07, de todas as unidades que compunham cada setor censitário, com registros das unidades não-residenciais – lojas, igrejas etc. – e de todos os

domicílios coletivos e particulares, classificando os particulares em ocupados, fechados, vagos ou de uso ocasional. Para cada domicílio coletivo encontrado, foram enumeradas, no CD 1.03, as famílias e as pessoas sós nele residentes.

As entrevistas eram realizadas simultaneamente ao processo de listagem, a partir da seleção da amostra, que determinava os domicílios particulares e, nos domicílios coletivos, as famílias e pessoas sós, onde seriam utilizados o CD 1.01 ou o CD 1.02.

Atribuições do Supervisor

Todas os procedimentos do trabalho do Supervisor estavam detalhados no CD 1.10. Suas atribuições, durante o Censo Experimental, foram as seguintes:

antes do início do trabalho de coleta

- fazer o reconhecimento das áreas de trabalho, percorrendo os setores sob sua responsabilidade, para rever e atualizar, nos CD 1.06, os mapas e as descrições e numerar as quadras e as faces dos quarteirões;

- revisar e atualizar a versão preliminar do Cadastro de Segmentos de Logradouros para os setores urbanos e para os aglomerados rurais de extensão urbana dos municípios com mais de 25 000 habitantes, segundo a Contagem da População 1996; e

- verificar a disponibilidade, em quantidades suficientes, de todo o material necessário à realização do trabalho de coleta pela sua equipe de Recenseadores.

durante o trabalho de coleta

- organizar a agenda de visitas semanais dos Recenseadores ao posto de coleta para entrega de material coletado;

- orientar sua equipe quanto à localização dos setores;

- retreinar os Recenseadores;

- acompanhar, em campo, cada um dos Recenseadores durante a primeira semana de coleta;

- fazer a crítica semanal do material coletado, durante todo o período de coleta do setor, obedecendo à rotina de crítica estabelecida, cujos pontos principais eram: crítica visual dos questionários e das Folhas de Coleta preenchidas, realizada no posto de coleta, na presença do Recenseador; reentrevistas em quatro domicílios selecionados, segundo os critérios da rotina; e verificação das possíveis omissões ou inclusões indevidas de unidades nos trechos percorridos pelo Supervisor, correspondentes aos domicílios selecionados para reentrevistas;

- acompanhar as situações de recusa às entrevistas, auxiliando os Recenseadores na tarefa de convencimento dos moradores, tendo em vista a prestação das informações;

- observar, nas Folhas de Coleta, os registros de domicílios fechados, alertando o Recenseador para o retorno periódico, com o fim de realizar a coleta das informações; e

- manter sempre atualizados os registros na Agenda do Supervisor, fazendo a anotação das ocorrências – erros, divergências nas reentrevistas, decisões e orientações – durante a aplicação da rotina de crítica e da produção semanal dos Recenseadores – número de questionários e de pessoas entrevistadas – para lançamento no Sistema de Informações Gerenciais da Coleta.

após o término da coleta

- verificar a transcrição dos totais das Folhas de Coleta para a Caderneta do Setor, feitos pelo Recenseador;

- atualizar o Cadastro de Segmento de Logradouros, com o total de unidades visitadas e de domicílios particulares ocupados;

- fazer o empastamento dos questionários, providenciando sua remessa, através de guia própria, à coordenação da UR;

- aguardar a autorização do ACM para o empastamento e remessa das Folhas de Coleta; e

- encaminhar ao ACM o restante do material preenchido – cadernetas dos setores, folhas de domicílio coletivo e a agenda do supervisor – e o excedente que não foi utilizado.

Atribuições do Agente Censitário Municipal

Coube ao ACM fazer o elo entre os Coordenadores e os Supervisores, estando, diretamente subordinado ao Coordenador de Subárea. No Censo Experimental, foi o responsável pela instalação e organização do posto de coleta, sendo gerente de um grupo de Supervisores. Suas atribuições foram:

- conhecer, previamente, os setores que compunham a área de pesquisa do posto de coleta;

- definir, em conjunto com o Coordenador de Subárea, as áreas de atuação de cada Supervisor;

- coordenar as atividades de reconhecimento prévio dos setores pelos Supervisores e Recenseadores;

- participar, como instrutor, do treinamento dos Supervisores e Recenseadores;

- auxiliar os Supervisores, buscando demover os moradores de eventuais recusas na prestação de informações ao Censo;

- zelar pelo bom padrão de qualidade das informações coletadas e pelo cumprimento dos prazos estabelecidos para a coleta; e

- fazer, quando responsável por um posto de coleta informatizado, a entrada semanal de dados no Sistema de Indicadores Gerenciais da Coleta, para o acompanhamento e gerenciamento do trabalho de campo.

Atribuição dos Coordenadores

As instâncias de coordenação, definidas para o Censo Experimental, na Unidade Regional, foram as Coordenações Técnica e de Subárea. Ao Coordenador Técnico coube a responsabilidade de conduzir, técnica e administrativamente, o Censo na Unidade Regional; ao Coordenador de Subárea foi atribuído o mesmo tipo de tarefa, porém, exercida numa área menor, o município.

As tarefas do Coordenador Técnico foram as seguintes:

- dimensionar a quantidade necessária de postos de coleta a serem instalados na UR;

- coordenar o processo seletivo para a contratação do pessoal de coleta – ACMs, Supervisores e Recenseadores;

- supervisionar o processo de treinamento das equipes de campo;

- definir, com o Coordenador de Subárea, um esquema de supervisão alternativo para o trabalho de coleta dos setores rurais, não localizados na sede do município;

- controlar a quantidade e providenciar a distribuição de todo o material necessário ao andamento do trabalho de campo; e

- encaminhar as pastas dos questionários e das Folhas de Coleta ao Centro de Captura de Dados, no Rio de Janeiro.

O trabalho do Coordenador de Subárea correspondia à execução das seguintes tarefas:

- instalar e presidir a Comissão Municipal Censitária;

- decidir sobre as alterações ou atualizações nos mapas censitários;

- definir, com os ACMs, as áreas de atuação dos supervisores;

- fazer o treinamento dos ACMs, Supervisores e Recenseadores;

- acompanhar, realizando visitas periódicas, o trabalho de coleta nos postos sob sua jurisdição;

- fazer a atualização das informações do Sistema de Administração de Pessoal e do Sistema de Indicadores Gerenciais da Coleta, correspondentes aos postos não informatizados sob sua responsabilidade;

- acompanhar a implementação das providências, tendo em vista investigar e corrigir os problemas apontados pelos relatórios dos sistemas de acompanhamento e gerenciamento técnico e administrativo do trabalho de campo; e

- fazer o controle e encaminhamento das pastas de questionários e de Folhas de Coleta à sede da UR.

5.4.4.4 Características do entorno dos domicílios

O Censo Experimental foi, também, utilizado para testar os quesitos de caracterização do entorno dos domicílios, ou seja: "existência de identificação", "existência de iluminação pública" e "existência de calçamento/pavimentação" que, pela primeira vez, seriam incluídos na coleta do Censo Demográfico 2000.

O teste foi realizado no Município de Marília, tendo como objetivo a observação e o acompanhamento do trabalho do Recenseador e do Supervisor, em relação ao preenchimento desses quesitos, a saber:

-incluídos na Caderneta do Setor, para caracterização do setor como um todo; e

- incluídos na Folha de Coleta[2], para a caracterização dos logradouros do setor. Em especial, buscava-se identificar as dificuldades para a coleta desse tipo de informação, assim como a adequação das orientações contidas no Manual do Recenseador.

5.4.4.4 a Acompanhamento do trabalho de campo

Em relação às características do entorno dos domicílios, o acompanhamento do trabalho de campo constou de duas partes: observação, em campo, do trabalho do Recenseador e investigação dos resultados, realizada no posto de coleta.

Observação do trabalho do Recenseador durante a coleta

O acompanhamento do Recenseador, durante a coleta do Censo Experimental, estava previsto na rotina de crítica implementada pelo Supervisor e no trabalho dos Observadores durante a atividade de coleta (CENSO EXPERIMENTAL 1999: ORIENTAÇÕES AO OBSERVADOR, 1999).

Desta forma, setores de situação rural e urbana, dos tipos não especial e especial de aglomerado subnormal[3], foram selecionados para serem visitados. durante o acompanhamento do trabalho de campo.

A observação do preenchimento das Folhas de Coleta quanto aos três quesitos de caracterização do entorno, foi o principal aspecto considerado. Nesse acompanhamento, as diversas situações encontradas nas áreas visitadas foram registradas por fotografias (COSTA, 1999b), ilustrando os principais aspectos observados.

Investigação das informações coletadas

Essa etapa consistiu na avaliação direta do material disponível, realizada no posto de coleta, que se constituiu na base de operação[4] deste trabalho. Inicialmente, o material selecionado para avaliação correspondeu às Cadernetas dos Setores e Folhas de Coleta dos setores censitários visitados. Mais tarde, foi selecionado e avaliado o material de outros setores, que não haviam sido objeto

[2] Constituem instrumentos de coleta e constam como anexos neste capítulo.
[3] Ver quadro de situação e tipo do setor no capítulo 7.
[4] Corresponde ao posto de coleta número 3. Havia três postos instalados em Marília durante a atividade censitária.

de observação em campo. A relação dos setores visitados e avaliados em gabinete encontra-se mais detalhada em Costa (1999a).

Esse trabalho de investigação dos dados permitiu, também, observar a atuação do Supervisor, tendo em vista que era sua tarefa verificar o preenchimento dos três quesitos na Folha de Coleta e na Caderneta do Setor, realizados pelo Recenseador.

5.4.4.4 b Avaliação do trabalho de campo

Os aspectos seguintes, referentes ao resultado da avaliação da investigação das características do entorno dos domicílios, no Censo Experimental, merecem ser destacados.

Acompanhamento do Recenseador

A maior dificuldade observada no trabalho dos Recensedores quanto ao preenchimento dos quesitos, foi a caracterização nas Folhas de Coleta dos setores de situação rural e de aglomerado subnormal. Este fato pode estar associado, nestas áreas, à maior irregularidade no traçado dos logradouros e à falta de quarteirões.

As orientações contidas no manual foram mais adequadas para a caracterização do entorno, nos setores urbanos do tipo não especial que apresentam estrutura urbana mais definida. Isto pode ser motivo para que os resultados tendam a apresentar-se mais consistentes para esses setores.

Entre os três quesitos investigados, o referente ao "calçamento/ pavimentação" foi o que apresentou menor dificuldade no preenchimento da Folha de Coleta. Isto pode ser explicado pelo fato dessa característica ser de mais fácil mensuração, por envolver menor grau de subjetividade. Por outro lado, o quesito referente à "identificação" foi o que envolveu maior subjetividade e, portanto, maior dificuldade para caracterização, sobretudo, nos setores de aglomerados subnormais e nos setores rurais. Isto pode ser atribuído à falta de intervenção pública nestes espaços, ao padrão irregular do traçado dos logradouros, assim como à extensão dos logradouros nos setores rurais.

Diferença nos registros na Caderneta do Setor e nas Folhas de Coleta

De forma geral, a adequação dos procedimentos do manual para o preenchimento desses dois instrumentos foi diferenciada; os critérios foram mais adequados e apresentaram maior facilidade à caracterização na Folha de Coleta. No caso da Caderneta do Setor, aconteceu o contrário, ocorrendo, inclusive, a omissão de marcação destes campos. Identificou-se outro problema quando, em alguns casos, os preenchimentos na caderneta não estavam compatíveis com os registros das Folhas de Coleta.

Para o Censo 2000, a partir do que foi observado em Marília, houve proposta de alteração na orientação para o preenchimento do quesito "existência de iluminação pública". A proposta de mudança considerava a existência de iluminação, quando em qualquer dos lados, na face ou no trecho do logradouro percorrido[5], houvesse ponto fixo – poste – de iluminação, uma vez que nas instruções para o Censo Experimental, era considerado, apenas, o ponto de luz no lado percorrido pelo Recenseador.

[5] Ver Manual do Recenseador - CD 1.09, p.50.

Foi verificada, também, a necessidade de atenção dos instrutores do treinamento do Censo 2000, de modo a reforçar a orientação para o preenchimento dos quesitos de caracterização do entorno dos domicílios.

5.4.4.5 Sistema de Indicadores Gerenciais da Coleta

O Sistema de Indicadores Gerenciais da Coleta – SIGC – teve como objetivo gerar informações que permitissem às diversas instâncias gerenciais do Censo Experimental, na sede do IBGE e nas Unidades Regionais de São Paulo e Pará, acompanhar a evolução da coleta e os níveis de cobertura, além de fornecer dados para o Sistema de Administração de Pessoal, com vistas à liberação do pagamento dos Recenseadores.

Descrição do Sistema

O acesso ao SIGC era realizado, via computador, em um posto de coleta informatizado, através da conexão à Internet. Ao ser acessado, o sistema exigia do usuário a digitação de senha para, em seguida, apresentar as seguintes opções na tela de abertura: Atualização, Consulta, Setores Liberados, Setores Pendentes e Consulta Nível Brasil.

Na opção Atualização, os setores de Marília e Bonito eram colocados em trabalho – status em andamento – e era feita a entrada de dados referente ao andamento e à conclusão da coleta.

Na opção Consulta, eram obtidas as informações do quantitativo de setores, segundo os diversos status de coleta, ou seja: não iniciados, em andamento, concluídos, liberados e paralisados.

A identificação dos setores já liberados para pagamento podia ser consultada através da opção Setores Liberados, enquanto a relação daqueles, cujos dados se afastavam de padrões estabelecidos por um conjunto de indicadores, era acessada através da opção Setores Pendentes.

Todas as opções ofereciam oportunidade de trabalho das informações, para o Municípios – Marília e Bonito – e para seus distritos, subdistritos e setores.

A opção Consulta Nível Brasil permitia configurar o SIGC, através da introdução das matrículas dos Recenseadores e da criação dos postos de coleta. Essa opção era de acesso restrito ao Coordenador Técnico e ao Coordenador de Subárea.

Atualização das informações

O SIGC recebia informações do Censo, segundo dois seguimentos básicos:

- recepção de informações parciais, permitindo o gerenciamento da coleta semanal do setor censitário, enquanto estivesse em trabalho; e

- recepção de informações, quando do encerramento da coleta, que levavam à avaliação final do trabalho realizado pelo Recenseador e ao seu pagamento.

Para cada setor com o status em andamento, após a entrega semanal do material coletado, eram transcritos para o SIGC o número de homens e mulheres recenseados e a quantidade de Questionários Básicos e da Amostra utilizados nas entrevistas. Essas informações eram retiradas da Folha de Acompanhamento Semanal da Coleta, que tinha como fonte os lançamentos da Caderneta do Supervisor.

Para os setores onde o Recenseador dava como encerrado o trabalho de coleta, eram repassadas ao sistema os dados do resumo, constantes da capa da Caderneta do Setor.

Sistema de indicadores

Tendo em vista alguns problemas detectados durante o acompanhamento do trabalho de campo nos censos anteriores, foram incorporados ao SIGC indicadores que permitiam avaliar se as informações referentes à coleta dos setores estavam em conformidade com padrões esperados, estabelecidos estatisticamente. Esses padrões foram calculados com base em informações passadas, relativas ao Censo 1991 e à Contagem da População 1996.

O padrão esperado para cada indicador era determinado, separadamente, para os setores urbanos e rurais, através de limites, inferior e superior, calculados imediatamente após à atualização semanal dos dados do setor.

Para os setores com coleta em andamento, foram definidos três indicadores, a saber:

- indicador A1 – total de pessoas recenseadas por dia de coleta:

$$\mathrm{A1}_i = \frac{PA_i + PNA_i}{X_i} \quad i = 1,2,3,....$$

onde: *i* identifica cada um dos setores;

PA_i é o número de pessoas recenseadas nos questionários da amostra;

PNA_i é o número de pessoas recenseadas nos questionários da não amostra; e

X_i é o número de dias de coleta.

- indicador A2 – número médio de pessoas recenseadas por questionário da amostra:

$$\mathrm{A2}_i = \frac{PA_i}{L_i} \quad i = 1,2,3,...$$

sendo L_i o número de questionários da amostra; e

- indicador A3 – razão entre o número de questionários da amostra e o total de questionários:

$$\mathrm{A3}_i = \frac{L_i}{(K_i + L_i)} \quad i = 1,2,3,....$$

onde K_i é o número de questionários da não amostra.

Para os setores concluídos, o SIGC calculava os indicadores C1 e C2, iguais, respectivamente, a A2 e A3, definidos anteriormente, acrescidos dos seguintes:

- indicador C3 – número médio de pessoas recenseadas por domicílio particular ocupado:

$$C3_i = \frac{PDA_i + PDNA_i}{(D_i + C_i)} \quad i = 1,2,3,....$$

PDA_i é o número de pessoas recenseadas nos domicílios particulares da amostra;

$PDNA_i$ é o número de pessoas recenseadas nos domicílios particulares da não amostra; e

D_i e C_i são, respectivamente, os números de domicílios particulares ocupados da amostra e da não amostra;

- indicador C4 – número de domicílios particulares ocupados; somente foi possível calcular esse indicador para os setores que se mantiveram inalterados entre 1991 e 1999; ou seja, não foram, nesse período, divididos ou grupados a outros setores, bem como não tiveram alterações em seus limites:

$$C4_i = (D_i + C_i)$$

- indicador C5 – número de unidades visitadas; como na situação anterior, só foi possível o cálculo desse indicador para os setores que permaneceram inalterados entre 1991 e 1999:

$$C5 = UNR_i + UO_i + F_i + V_i + D_i + C_i$$

onde UNR_i, UO_i, F_i e V_i são, respectivamente, o número de unidades não residenciais, de domicílios de uso ocasional, fechados e vagos;

- indicador C6 – razão entre o número de domicílios fechados e vagos e o total de domicílios particulares ocupados, fechados e vagos:

$$C6_i = \frac{F_i + V_i}{(F_i + V_i + D_i + C_i)} \quad i = 1,2,3,....$$

- indicador C7 – razão entre o número de domicílios de uso ocasional e o total de domicílios particulares ocupados e de uso ocasional:

$$C7_i = \frac{UO_i}{(UO_i + D_i + C_i)} \quad i = 1,2,3,....$$

- indicador C8 – razão de sexo para os domicílios particulares:

$$C8_i = \frac{H_i}{(H_i + M_i)} \quad i = 1,2,3,....$$

onde H_i e M_i são, respectivamente, o número de homens e de mulheres recenseados nos domicílios particulares.

Os limites de variação, inferior – LI – e superior – LS – foram estabelecidos para os indicadores da seguinte forma:

- indicador A1 – fixados a partir da observação de informações sobre a coleta do Censo 1991, como sendo: setores urbanos LI = 12 e LS = 48; e setores rurais LI = 9 e LS = 36;

- indicadores C4 e C5 – calculados, tendo como base dados do Censo de 1991 e Contagem da População 1996, a partir da razão;

$R_{i\ (1996/1991)} = C_{i(1996)} / C_{i(1991)}$, onde C = C4 ou C5,

corrigida para o período de 3 anos – de 1996 a 1999 – chegando-se a

$R_{i\ (1996/1999)} = (C_{i(1996)} / C_{i(1991)})3^{3/5}$;

em seguida, a partir das distribuições de $R_{(1996/1999)}$, obteve-se os limites inferiores LI e superiores LS dessas razões de crescimento como

$LI\ [R_{(1996/1999)}] = Q25[R_{(1996/1999)}] - 2{,}5 \times (Q75[R_{(1996/19999)}] - Q25[R_{(1996/1999)}]$ e

$LS[R_{(1996/1999)}] = Q75[R_{(1996/1999)}] + 2{,}5 \times (Q75[R_{(1996/1999)}] - Q25[R_{(1996/1999)}]$;

onde Q25 e Q75 são, respectivamente, o primeiro e terceiro quartis das distribuições da razão R, calculados para os setores concluídos, da agência de coleta – Marília ou Capanema - ou da Unidade da Federação – São Paulo ou Pará – quando o número de observações para as agências era considerado insuficiente.

Por fim, chegou-se aos seguintes limites para os indicadores:

$LI = LI\ [R_{(1996/1999}] \times C_{(1996)}$ e

$LS = LS\ [R_{(1996/1999}] \times C_{(1996)}$

- demais indicadores – calculados, tendo como referências as informações do Censo 1991, como:

$LI = Q25[I_{(1991)}] - 2{,}5 \times (Q75[I_{(1991)}] - Q25[I_{(1991)}])$ e

$LS = Q75[I_{(1991)}] + 2{,}5 \times (Q75[I_{(1991)}] - Q25[I_{(1991)}])$,

assumindo-se I como cada um dos indicadores e tendo em conta que, para C6 e C7, o valor de LI foi igual a zero.

À medida que os valores observados de um ou mais indicadores encontravam-se fora dos limites LI e LS, automaticamente, o sistema assumia a condição de pendente para o setor em avaliação. Nessa oportunidade, eram emitidas mensagens de advertência – valor do indicador acima ou abaixo do esperado – que levavam à tomada de providências, tendo em vista identificar e corrigir os possíveis equívocos. Tratando-se de setor com coleta concluída, o respectivo pagamento ao Recenseador só era autorizado, quando as pendências fossem, totalmente, esclarecidas.

Uma descrição detalhada das ações gerenciais, que eram implementadas em atenção às mensagens de advertência dos indicadores, pode ser encontrada no documento *Censo experimental 1999: instruções operacionais para agentes censitários municipais e coordenadores do censo experimental* (1999).

A investigação das situações de pendência passava, em primeiro lugar, pela revisão de todas as transcrições e somas que determinavam os valores lançados no SIGC. Caso não fossem encontrados erros, acontecia o retorno a campo do Supervisor e/ou Recenseador para verificação das situações que poderiam ter determinado a condição ou condições de pendência.

No caso de setor com coleta em andamento, as possíveis modificações dos dados decorrentes do processo de investigação podiam realimentar o SIGC de imediato, ou aguardarem a próxima atualização semanal, para que o setor, novamente, viesse a ser submetido ao crivo dos indicadores.

Por outro lado, se o setor pendente já tivesse a coleta concluída, os novos dados seriam, imediatamente, inseridos no SIGC, para que uma nova avaliação fosse, de pronto, elaborada, objetivando liberar o setor para pagamento o mais rápido possível. Caso a investigação determinasse a manutenção dos valores originalmente lançados, o setor continuaria com o status de pendente e somente o Coordenador Técnico ou o Coordenador de Subárea poderiam autorizar o seu pagamento, após receberem justificativas convincentes do ACM.

Alterações no Sistema

Feitas as avaliações do SIGC, foram recomendadas as seguintes alterações, tendo em vista a sua melhor utilização no Censo 2000:

- considerando o elevado número de erros observados durante a fase de entrada de dados, decidiu-se pela ampliação do conjunto de críticas que atuavam nessa fase, sobretudo, as que verificavam inconsistências entre as datas de início da coleta, das atualizações das informações semanais e do encerramento da coleta; e

- objetivando diminuir o tempo de conexão ao sistema, recomendou-se a criação de uma variante na opção Atualização, de modo que o operador pudesse fazer, *off line*, os lançamentos das informações setor a setor, conectando-se ao SIGC, somente quando do efetivo envio, em bloco, das atualizações de dados ao sistema.

5.4.5 Pesquisa de Avaliação da Cobertura da Coleta

A Pesquisa de Avaliação da Cobertura da Coleta do Censo Experimental – PA - teve como objetivo testar as fases de coleta, de batimento e de reconciliação das informações, incluídos todos os instrumentos utilizados, os procedimentos estabelecidos, o sistema de entrada de dados descentralizado e o treinamento das equipes da pesquisa.

O planejamento da PA foi realizado paralelamente ao do Censo Experimental e pressupôs a utilização de amostragem, com a seleção de setores, não especiais e de aglomerados subnormais, e de domicílios.

O trabalho teve início com a coleta das informações e prosseguiu com as fases de: batimento – confronto de informações – e de reconciliação – retorno aos domicílios para averiguação das informações consideradas divergentes.

O erro de cobertura do Censo é uma conseqüência direta da omissão ou inclusão indevida de domicílios particulares e das pessoas neles residentes, assim como das pessoas residentes em domicílios particulares ocupados e considerados os mesmos no Censo e na PA.

A medição do erro de cobertura é feita a partir da construção de indicadores de omissão de domicílios e pessoas. O método escolhido para estimação desses indicadores foi o *Dual System Estimation*, o que pressupunha que a coleta da PA fosse realizada nos mesmos moldes que a do Censo Experimental e que houvesse total independência entre elas. Essa independência deve acontecer em relação ao manuseio do material de coleta das duas pesquisas e ao elenco de Recenseadores e de Supervisores.

5.4.5.1 Amostra da Pesquisa

No tocante à definição do esquema de amostragem a ser utilizado para a PA, ficou decidida a utilização de um desenho em dois estágios de seleção – no primeiro eram selecionados setores e no segundo, os domicílios.

Com o primeiro estágio de seleção, o objetivo era estimar a omissão do Censo Experimental, em nível de domicílio; com o segundo estágio – amostra de domicílios – o intuito era avaliar a omissão de pessoas. De acordo com o estabelecido por Oliveira (1999), o tamanho da amostra de setores era obtido por:

$$n = \frac{N\delta^2\hat{\gamma}^2}{N(er)^2 + \delta^2\hat{\gamma}^2}$$

onde:

***N* é o total de setores existentes em Marília e Bonito;**

δ é o múltiplo do desvio padrão (1, considerando um grau de confiança de 68%);

er é o erro relativo de amostragem = 0,20;

$\hat{\gamma}$ é o coeficiente de variação da variável "número de domicílios omitidos", calculado com base nas informações da Pesquisa de Avaliação do Censo 1991, levando em conta o setor (conglomerado);

$\hat{\gamma} = \frac{s_e}{\bar{y}}$, sendo que: $s_e^2 = \frac{1}{n^* - 1}\sum_{i=1}^{n^*}(y_i - \bar{y})^2$ e $\bar{y} = \frac{\sum_{i=1}^{n^*} y_i}{n^*}$

n^* é o número de setores selecionados para a PA do Censo 1991 para São Paulo; e

y_i é o número de domicílios omitidos com código de batimento igual a 171 – domicílio particular ocupado na PA e omitido no Censo, cujos moradores nele residiam na data do Censo – e 172 – domicílio particular ocupado na PA e omitido no Censo, cujos moradores passaram a residir após a data do Censo – no i-ésimo setor selecionado na PA do Censo 1991.

Para a seleção dos domicílios, foi utilizada uma fração de amostragem de 10%, sendo essas unidades selecionadas dentro dos setores escolhidos para comporem a amostra da PA.

O resultado apontou que, dos 300 setores de Marília e 12 de Bonito, fossem selecionados, respectivamente, 53 e 9, para comporem a amostra de primeiro estágio da amostra da PA.

5.4.5.2 Instrumentos de coleta, de batimento e reconciliação

Os instrumentos utilizados na pesquisa foram os seguintes:

PA 1.01 – Questionário da Amostra – aplicado em todos os domicílios particulares selecionados para amostra que estavam ocupados no dia da visita do Recenseador. Destinava-se ao registro de informações referentes aos domicílios, às pessoas moradoras, na data da PA, e àquelas que neles residiam, em 31 de julho de 1999, e morreram ou mudaram de domicílio;

PA 1.06 – Caderneta do Setor – documento que, além de apresentar o mapa e a descrição do setor, servia para o acondicionamento das Folhas de Coleta utilizadas;

PA 1.07 – Folha de Coleta – formulário destinado ao registro de todas as unidades residenciais e não residenciais pertencentes ao setor;

PA 1.09 – Manual do Recenseador – continha todos os conceitos da pesquisa e as instruções necessárias ao preenchimento dos questionários, da Folha de Coleta e da Caderneta do Setor;

PA 1.10 – Manual de Batimento e Reconciliação – apresentava as instruções necessárias ao trabalho de batimento e de reconciliação das informações coletadas;

PA 1.11 – Manual de Acompanhamento e Empastamento – documento que trazia orientações para o trabalho de supervisão da coleta e para o empastamento de todo o material coletado; e

PA 1.12 – Manual de Entrada de Dados – mostrava as orientações para a entrada dos dados consolidados, após o batimento e a reconciliação das informações.

5.4.5.3 Coleta das informações

O treinamento para a fase de coleta foi realizado nas Agências de Marília e Capanema[6] e teve a duração de apenas um dia, visto que os Recenseadores e Supervisores da pesquisa foram selecionados entre os que haviam atuado no Censo Experimental.

O objetivo desse treinamento foi repassar com os treinandos todos os conceitos, sanando possíveis dúvidas, enfatizando a correta aplicação da fração de amostragem nos setores, e apresentar e explicar o preenchimento dos

[6] O treinamento da coleta, o treinamento do batimento e da reconciliação e os trabalhos correspondentes foram realizados em Capanema devido à falta de infra-estrutura no Município de Bonito.

instrumentos de coleta da PA, a Caderneta do Setor, a Folha de Coleta e o Questionário da Amostra.

O início efetivo da coleta, no Município de Bonito, ocorreu no dia 11 de outubro, e no Município de Marília, a partir do dia 5 desse mesmo mês. O tempo gasto nesse trabalho foi, aproximadamente, um mês, no Município de Marília e dez dias, em Bonito.

Uma vez que a coleta das informações e a supervisão da PA foram realizadas por Supervisores que atuaram na coleta do Censo Experimental, ficou decidido que não poderiam trabalhar nos mesmos setores anteriores. Utilizaram-se 16 Recenseadores e 2 Supervisores, em Marília, e um Recenseador e um Supervisor, em Bonito.

Devido a restrições orçamentárias, dos 53 setores planejados para a amostra de Marília, somente 16 puderam ser selecionados, sendo 12 urbanos e quatro rurais; já em Bonito, dos nove previstos, apenas um setor urbano foi selecionado. Foram listados, respectivamente, 2 544 e 240 domicílios particulares ocupados em Marília e Bonito.

O trabalho do Supervisor consistiu no seguinte:

- identificar, em campo, os setores selecionados;

- orientar os Recenseadores para reconhecimento dos setores;

- verificar o preenchimento dos instrumentos de coleta;

- realizar reentrevistas; e

- informar, semanalmente, aos coordenadores o andamento dos trabalhos de coleta.

Quanto ao Coordenador de Subárea, era de sua competência montar as Cadernetas dos Setores selecionados, orientar e sanar as dúvidas dos Supervisores, providenciar a distribuição do material, informar à CTD sobre o andamento dos trabalhos e fornecer o quantitativo de domicílios particulares arrolados e de domicílios particulares selecionados, nos setores liberados.

5.4.5.4 Batimento e reconciliação das informações

Para essa fase do trabalho da PA, o treinamento foi ministrado, em Marília e Capanema, por dois instrutores, em cada uma dessas agências.

Anteriormente ao cálculo dos indicadores de omissão de domicílios e de pessoas, é necessária a execução do Batimento das Informações coletadas pelas folhas de coleta e pelos questionários da PA e do Censo, o que corresponde à parte da apuração da PA. Como durante essa fase costumam ocorrer casos de dúvida quanto a possível omissão de domicílios e/ou pessoas, para melhor avaliar a precisão da coleta do Censo, exige-se um retorno aos domicílios para averiguação; a esse processo dá-se o nome de Reconciliação das Informações.

O Batimento das Informações foi realizado, inicialmente, através do confronto dos registros nas Folhas de Coleta das duas pesquisas. Nesse confronto, a cada unidade registrada foi atribuído um código, segundo os critérios estabelecidos no Manual de Batimento e Reconciliação – PA 1.10. Por exemplo: o código 410 indicava tratar-se de um domicílio particular vago na PA e domicílio particular ocupado no Censo; e o código 171 correspondia a um domicílio particular ocupado na PA, que não foi encontrado no Censo.

O Manual de Batimento e Reconciliação continha orientações sobre o reconhecimento dos setores e orientava a procura das unidades registradas na PA e não encontradas no Censo, também, nos setores vizinhos.

Após a conclusão do confronto das informações das Folhas de Coleta, dava-se seqüência à fase de batimento, com a comparação das informações dos questionários correspondentes aos domicílios particulares ocupados na PA, considerados os mesmos do Censo. Para isso, eram utilizadas informações como "nome", "sexo", "parentesco", "idade", "se residia ou não no domicílio na data do censo", "ano em que passou a residir no domicílio" etc.

Foram aplicados códigos às pessoas registradas nos questionários das duas pesquisas, de acordo com os critérios estabelecidos no PA 1.10, de modo que retratassem a situação de moradia de cada pessoa. A consolidação dos códigos e, conseqüentemente, dos quantitativos foram lançados no PA 1.14 e PA 1.15, correspondendo, respectivamente, à alocação de pessoas por código e por faixa etária, em cada domicílio confrontado com os códigos 111, 115,e 117. A partir daí, avalia-se a omissão de pessoas por faixa etária, em domicílios particulares ocupados, considerados os mesmos no Censo e na PA e, ao final, a omissão global de pessoas.

Por outro lado, a fase de reconciliação consistiu na checagem das informações duvidosas decorrentes do trabalho na fase de batimento. Optou-se pela reconciliação dos códigos atribuídos que geravam possíveis omissões de domicílios ocupados no Censo e na PA. Este trabalho representou o retorno a campo nos domicílios onde foram atribuídos tais códigos, sendo realizado pelos mesmos técnicos que participaram do trabalho de batimento.

Em Marília, os trabalhos de batimento e reconciliação foram realizados, simultaneamente, por quatro técnicos, sendo dois contratados. Esse pequeno número fez com que essas tarefas se estendessem por, aproximadamente, 50 dias. Em Capanema, dois técnicos do IBGE encerraram o batimento e a reconciliação das informações em quatro e dois dias, respectivamente.

5.4.5.5 Acompanhamento da coleta, batimento e reconciliação

Um outro objetivo, a ser alcançado na PA, foi testar uma forma de comunicação entre a sede do IBGE e as Unidades Regionais, através da qual fosse possível a obtenção de informações sobre o andamento dos trabalhos desenvolvidos, visando a sua implantação futura durante a PA do Censo 2000.

Com essa finalidade, foi desenvolvido um sistema de acompanhamento, utilizando Lotus Notes e tecnologia WEB-Intranet. Assim, os técnicos responsáveis pela PA, na DPE, eram informados, durante a fase de coleta, sobre o número total de domicílios particulares e o número de domicílios particulares selecionados em cada setor.

Informações sobre o andamento semanal dos trabalhos, em cada uma das fases de coleta, batimento e reconciliação das informações, estavam, também, disponíveis, correspondendo ao total de setores com trabalho não iniciado, em andamento, interrompido ou liberado. Todos os aspectos ligados ao acompanhamento dos trabalhos da PA estão descritas no Manual de Acompanhamento e Empastamento – PA 1.11.

5.4.5.6 Entrada de dados

O trabalho de entrada de dados era realizado somente ao final das fases de batimento e reconciliação das informações de cada setor. Nessa oportunidade, era utilizado o *Integrated Microcomputer Processing System* – IMPS através

da execução do módulo *Census and Survey Data Entry* – CENTRY que oferecia um conjunto de planilhas especialmente elaboradas para o trabalho. Simultaneamente, fazia-se uso, para identificação de erros e inconsistências, de um outro módulo, denominado *Consistency and Correction* – CONCOR.

As planilhas preenchidas eram enviadas, diretamente, à DPE, através do Lotus Notes. Após o recebimento dos dados do setor, era realizada uma investigação, tendo em vista identificar pendências ou inconsistências. Sanados todos os problemas, os dados estavam prontos para o cálculo das estimativas dos indicadores de omissão de domicílios e de pessoas.

5.4.5.7 Método de estimação

O método usado para o cálculo dos indicadores de omissão foi o *Dual System Estimation*[7]. Sua utilização requer independência na coleta das duas pesquisas e pressupõe o confronto das informações da seguinte maneira.

PA	Censo		
	Total	Unidades incluídas	Unidades não-incluídas
Total	**t**	**a + c**	**b + d**
Unidades incluídas	a + b	a	b
Unidades não-incluídas	c + d	c	d

onde:

a é o número de unidades incluídas no Censo e na PA;

b é o número de unidades incluídas apenas na PA;

c é o número de unidades incluídas apenas no Censo;

d é o número de unidades desconhecidas que não foram incluídas no Censo e na PA (desconhecido); e

t é o total de unidades da população.

Na hipótese de que, em cada setor da amostra, a primeira pesquisa realizada tenha sido o Censo e a segunda a PA, utilizando o estimador decorrente do *Dual System Estimation*, temos a taxa de omissão, obtida por

$$\hat{R} = \frac{\hat{b}}{(\hat{a} + \hat{b})}$$

Vale ressaltar que, devido ao fato de ter sido reduzido o tamanho da amostra de setores – de 53 para 16, em Marília, e de nove para um, em Bonito, – a precisão das estimativas dos indicadores ficou bastante prejudicada. Foram registrados como domicílios particulares ocupados em Marília, 2 337 no Censo e na PA 2544; em Bonito, esses números foram de 244, no Censo, e 240, na PA. Foram levados à fase de reconciliação 248 e 23 domicílios em Marília e Bonito, respectivamente.

[7] *Ver* Marks (1973).

A partir dos dados recebidos, foram calculadas as seguintes taxas de omissão, separadamente, com e sem o trabalho de reconciliação das informações:

- taxa de omissão de domicílios particulares ocupados;

- taxa global de omissão de pessoas em domicílios particulares ocupados;

- taxa de omissão de pessoas em domicílios particulares ocupados, considerados os mesmos nas duas pesquisas; e

- taxas de omissão de pessoas em domicílios particulares ocupados, considerados os mesmo nas duas pesquisas, segundo as seguintes faixas de idade: menos de um ano, um a quatro anos, 5 a 14 anos, 15 a 59 e 60 anos ou mais.

Devido aos problemas já citados, esses resultados não foram considerados como indicadores de omissão para o Censo Experimental.

5.4.5.8 Processo de Avaliação

O modelo de treinamento proposto foi considerado adequado para a capacitação das equipes da PA, no que toca à realização das suas tarefas. No entanto, ficou claro também que os futuros treinandos teriam bastante dificuldade em retransmitirem as instruções referentes às fases de Batimento e Reconciliação das Informações, o que apontava como inoportuna, a idéia de que o treinamento da PA do Censo 2000 fosse realizado através dessa alternativa de repasse de instrução.

Por outro lado, os trabalhos da PA foram avaliados pelos técnicos da equipe de campo, quando responderam às indagações apresentadas nos seguintes questionários:

- Questionário de Avaliação do Trabalho – Coleta de Dados – avaliava a apresentação, formatação, diagramação e facilidade de preenchimento dos questionários e da Folha de Coleta, bem como o conteúdo do Manual do Recenseador – PA 1.09, quanto à clareza de suas instruções; e

- Questionário de Avaliação do Trabalho – Batimento, Reconciliação e Entrada de Dados – avaliava a disponibilidade do material no período previsto para a realização das etapas, a facilidade para identificação das unidades e aplicação dos códigos e a compreensão das instruções do Manual de Batimento e Reconciliação – PA 1.10 e do Manual de Entrada de Dados – PA 1.12.

Ao final de cada questionário, havia um espaço destinado aos comentários pessoais de cada participante/avaliador do trabalho da PA. As avaliações mostraram a necessidade de pequenos ajustes nos instrumentos de coleta e um maior cuidado na instrução referente à abordagem dos informantes, ao preenchimento dos blocos quatro e cinco do questionário e ao lançamento das informações obtidas após a fase de reconciliação.

5.4.6 Teste do Projeto Vamos Contar

No Censo Experimental, foi, também, testado o projeto Vamos Contar!. No teste, foram envolvidas todas as escolas do Ensino Fundamental dos Municípios de Marília (SP) e de Bonito (PA).

Tabela 5.3 - Escolas, professores e alunos envolvidos no teste do projeto Vamos Contar!

	Marília/São Paulo	Bonito/Pará
Escolas	56	46
Professores	1 105	112
Alunos	31 246	3 142

Fonte: IBGE, Centro de Documentação e Disseminação de Informações, Projeto Vamos Contar!

As escolas receberam caixas, contendo o seguinte material:

- Cartas explicativas – informações sobre o projeto, direcionadas aos diretores e professores, aos pais dos alunos e aos integrantes do Conselho Escola/Comunidade;

- Folheto – com as principais informações sobre o Censo 2000;

- Cartazes – para divulgar o projeto nas escolas, estimulando a participação de todos;

- Guia do professor – propostas de planos de aula e de atividades para 1ª a 4ª série e 5ª a 8ª série do Ensino Fundamental;

- Mapa do Brasil – para 1ª a 4ª série e 5ª a 8ª série, em formato 1,00 x 0,90m, para cada sala de aula;

- Carta-resposta – ficha de avaliação do projeto para diretores e professores;

- Questionário do professor/fase experimental – para uma amostra de 4% dos professores de 1ª a 4ª série e 7% para os de 5ª a 8ª, buscando informações mais detalhadas sobre os resultados do uso do material; e

- Roteiro de entrevista com o aluno/fase experimental – para uma amostra de 0,76% dos alunos de 1ª a 4ª série e 0,87% para os de 5ª a 8ª, visando obter informações mais detalhadas sobre os resultados do uso do material.

Além dessas peças, foi elaborada Carta Explicativa aos Diretores sobre o plano de teste e orientações de procedimentos na aplicação do projeto.

A Coordenadora do Projeto Vamos Contar! e as pedagogas responsáveis pela elaboração do material foram aos dois municípios para avaliar a recepção e o uso do material nas escolas. Em algumas escolas, o projeto já estava sendo desenvolvido, ao passo que, em outras, o material ainda não havia sido distribuído aos professores, embora a escola o tivesse recebido.

Houve uma diferença na aplicação do projeto nos dois municípios. Em Marília, diretores e professores aproveitaram a presença da equipe do projeto para tirar dúvidas, solicitar orientação sobre o uso do Guia do Professor, sobre o questionário de avaliação e sobre outros itens relativos ao Vamos Contar! Essa experiência permitiu à equipe do projeto fazer avaliações sobre a distribuição do material, prazos e envolvimento das equipes das escolas, com vistas ao Censo 2000.

A maior parte das escolas utilizou o material. Os alunos tiveram aulas baseadas nos guias fornecidos pelo IBGE, desenvolvendo atividades ligadas, principalmente, à matemática, estudos sociais, geografia e português. Assim, além dos conteúdos das disciplinas que faziam parte dos parâmetros curriculares do MEC, os alunos tiveram a oportunidade de vivenciar questões ligadas à cidadania, motivadas pela explicação da importância do Censo Cemográfico, da participação da população e da necessidade de compreender e utilizar os resultados da pesquisa.

Já no Município de Bonito, foram encontradas algumas dificuldades: a escola estadual estava fechada para reformas, sem data para reabrir; a visita a algumas escolas mostrou que o trabalho com os alunos ainda não havia sido iniciado no final do mês de agosto; em várias escolas, o material não tinha sido distribuído adequadamente, ficando claro que as instruções não foram lidas; havia turmas de 5ª a 8ª série, funcionando em escolas municipais que não constavam da listagem do IBGE e que, portanto, não receberam material.

Diante desse quadro, a equipe do IBGE mudou a estratégia de trabalho. Com a ajuda do Diretor Estadual de Ensino de Bonito, reuniu-se com o máximo de professores e diretores para tentar motivá-los a utilizar o material com os alunos e avaliar o projeto. Mesmo assim, a resposta ao projeto em Bonito foi menor do que em Marília.

A equipe técnica responsável elaborou um relatório de testagem do material baseado no trabalho de observação de entrevistas, de aplicação de questionários, de palestras, de reuniões e de apuração das Cartas-respostas, enviadas pelos professores e diretores que participaram da experiência (SILVA; BARROS, 1999).

Após avaliação geral, o Projeto Vamos Contar! foi submetido ao Centro de Documentação e Disseminação de Informações e à apreciação do Conselho Diretor do IBGE, obtendo aprovação para sua implementação total, que contou com o envolvimento de cerca de 200 000 escolas em todo o País.

5.4.7 Pesquisa de Avaliação da Comunicação

A Pesquisa de Avaliação da Comunicação do Censo Experimental, realizada em setembro de 1999, com os moradores residentes em Marília, teve a finalidade de obter avaliações sobre os efeitos dos diversos meios de comunicação empregados no Censo Experimental. A realização da pesquisa permitiu a obtenção de estimativas de proporções de domicílios em que as pessoas que os representavam tomaram conhecimento do Censo Experimental mediante diferentes modos de divulgação. Para tanto foi aplicado um questionário específico de avaliação a uma amostra de domicílios. Essa pesquisa foi realizada no Município de Marília, após o término do trabalho de campo do Censo Experimental. O questionário utilizado encontra-se em anexo.

5.4.7.1 Planejamento da amostra

A população alvo da pesquisa era constituída por moradores de domicílios particulares permanentes, pertencentes a 251 setores urbanos e 15 setores rurais do Município de Marília. Foram excluídos da população alvo 16 setores selecionados para a Pesquisa de Avaliação do Censo Experimental, cinco setores especiais - cadeias e asilos – e 13 setores com menos de 50 domicílios. O sistema de referência utilizado para a pesquisa foi um arquivo de planejamento

da Base Operacional Geográfica 2000 que continha, para cada setor, a informação sobre a situação e o tipo do setor, além do número esperado de domicílios particulares ocupados e de unidades visitadas.

O desenho adotado para a amostra da pesquisa foi o de amostragem conglomerada em dois estágios, com estratificação das unidades primárias de amostragem. Os setores censitários formaram as unidades amostrais de 1º estágio - unidades primárias de amostragem (UPA) – e foram estratificados considerando a situação urbana ou rural. Os domicílios compuseram as unidades do 2º estágio da amostra – unidades secundárias de amostragem (USA).

O tamanho da amostra foi determinado considerando a interação entre os recursos operacionais disponíveis e o nível de precisão desejado para as estimativas. Vários cálculos baseados em amostragem aleatória simples foram efetuados a fim de definir o tamanho de amostra, fixando-se vários níveis de precisão para as estimativas de proporções. Para ajustar o tamanho de amostra obtido ao plano amostral utilizado na pesquisa foram considerados vários valores de efeito de plano amostral – EPA.

Decidiu-se por um tamanho de amostra necessário para estimar uma proporção pequena – 3% – com um CV de 25% e considerando um EPA de 2, resultando num total de cerca de 1 000 domicílios. O tamanho da amostra de setores ficaria entre 40 e 50 setores, variando o número de domicílios selecionados em cada setor entre 25 e 20.

Para o dimensionamento final da amostra de domicílios, considerou-se 25 Recenseadores para a pesquisa e um dia para a coleta dos questionários em cada setor, considerando-se que cada um faria a coleta em 2 setores; assim, decidiu-se por uma amostra de setores de tamanho 50. Houve um acréscimo do número de domicílios a serem selecionados em cada setor da amostra, de 20 para 25 unidades, em função da estratificação, que foi posteriormente definida, já que os cálculos foram feitos sem levar em conta a partição entre setores urbanos e rurais. O tamanho final da amostra ficou em 1 250 domicílios.

A alocação da amostra de setores, nos estratos considerados, foi feita por meio de alocação proporcional, com base no número de setores existentes em cada um dos estratos, setores urbanos e setores rurais.

5.4.7.2 Seleção da amostra de setores e de domicílios

A seleção dos 50 setores da amostra foi realizada, de forma sistemática, com probabilidade proporcional ao número de domicílios esperados em cada setor, com base no sistema de referência – o arquivo da Base Operacional Geográfica do Censo 2000. Uma avaliação quanto à distribuição dos setores selecionados nos três postos de coleta existentes no Município de Marília, não indicou a necessidade de ajustes na carga de trabalho dos postos.

A etapa de seleção da amostra de domicílios ocorreu, após o fechamento do trabalho de campo do Censo Experimental, naqueles setores selecionados. O número total de domicílios particulares permanentes, em cada setor, foi extraído do Sistema de Indicadores Gerenciais da Coleta – SIGC, implementado para o acompanhamento da coleta do Censo Experimental.

A seleção foi feita por amostragem aleatória simples, usando como cadastro a Folha de Coleta de cada setor selecionado. Apenas um setor apresentou uma situação bastante peculiar no que se refere ao número de domicílios particulares – apenas 2 domicílios particulares permanentes, 221 classificados

como unidades de uso ocasional e 28 como vagos. Por esse motivo, foi feita sua substituição, levando-se em conta para a seleção do novo setor, as variáveis de controle: posto de coleta, distrito, situação, tipo do setor e o número esperado de domicílios particulares ocupados - porque a seleção dos setores foi proporcional a esse número de domicílios.

5.4.7.3 Expansão da amostra

Como a pesquisa foi realizada por amostragem probabilística, cada domicílio selecionado na amostra representa, também, outros domicílios que fazem parte da população alvo. Assim, para cada unidade selecionada – domicílio -, foi associado um fator de expansão – peso - que representa, basicamente, o inverso da sua probabilidade de seleção.

$w_{hij} = \frac{1}{m_h}\frac{N_h}{N_{hi}}\frac{N^*_{hi}}{n^*_{hi}}$ (para todo domicílio j do i-ésimo setor do h-ésimo estrato);

onde:

m_h é o número de setores na amostra no h-ésimo estrato - urbano ou rural;

N_h é o total de domicílios particulares ocupados esperado em cada um dos 2 estratos;

N_{hi} é o número esperado de domicílios particulares ocupados no i-ésimo setor do h-ésimo estrato;

N^*_{hi} é o número de domicílios particulares permanentes no i-ésimo setor do h-ésimo estrato, obtido após a coleta do Censo Experimental; e

n^*_{hi} é número de domicílios particulares permanentes com entrevista realizada, na amostra da pesquisa, no i-ésimo setor do h-ésimo estrato, obtido após a coleta do Censo Experimental.

Para obter a estimativa de um total, multiplica-se o valor da característica de interesse pelo peso associado a cada registro de domicílio e somam-se os valores para todos os registros. As proporções são estimadas dividindo-se o número total estimado de unidades com uma determinada característica pelo número total estimado de unidades na população. Dessa forma, as estimativas de totais e proporções – porcentagens - usadas na avaliação dos resultados são calculadas, respectivamente, utilizando-se as seguintes expressões:

$$\hat{T} = \sum_{h,i,j} w_{hij} x_{hij} \qquad \hat{P} = \frac{\sum_{h,i,j} w_{hij}.x_{hij}}{\sum_{h,i,j} w_{hij}}$$

onde:

h varia nos estratos(1,2), i nos setores(1,...,) e j nos domicílios da amostra(1,...,);

$\hat{T}$ é o estimador de total das características de interesse da pesquisa;

$\hat{P}$ é o estimador de proporção das características de interesse da pesquisa;

x_{hij} é o valor da variável de interesse para o j-ésimo domicílio do i-ésimo setor do h-ésimo estrato; e

w_{hij} é o peso do j-ésimo domicílio do i-ésimo setor do h-ésimo estrato.

A avaliação da precisão das estimativas produzidas por pesquisas amostrais é um ponto fundamental do processo de produção de informações por amostragem. Dela depende o grau de confiança das conclusões analíticas advindas dos resultados dessas pesquisas. A precisão relativa das estimativas de interesse foi expressa como função do coeficiente de variação – CV – medido em porcentagem. Esses coeficientes de variação foram estimados, utilizando-se o Método do Conglomerado Primário (*Ultimate Cluster*), detalhado em Hansen e outros (1953) ou em Pessoa e Silva (1998), por meio do Sistema SUDAAN (*Survey Data Analysis*).

Apresenta-se, a seguir, o estimador do coeficiente de variação das estimativas das características de interesse da pesquisa.

$$\hat{CV}\left(\hat{T}\right)=\frac{\sqrt{\hat{V}\left(\hat{T}\right)}}{\hat{T}}$$

onde:

$$\hat{V}\left(\hat{T}\right)=\sum_{h}\frac{m_h}{m_h-1}\sum_{i}\left(\hat{T}_{hi}-\frac{\hat{T}_h}{m_h}\right)^2$$

$$\hat{T}_{hi}=\sum_{j}w_{hij}x_{hij}$$

$$\hat{T}_h=\sum_{i,j}w_{hij}x_{hij}$$

$\hat{V}\left(\hat{T}\right)$ é o estimador de variância das características de interesse da pesquisa;

$\hat{T}_{hi}$ é o estimador de total do i-ésimo setor do h-ésimo estrato;

$\hat{T}_h$ é o estimador de total do h-ésimo estrato; e

$\hat{T}$ é o estimador de total das características de interesse da pesquisa.

$$\hat{CV}\left(\hat{P}\right)=\frac{\sqrt{\hat{V}\left(\hat{P}\right)}}{\hat{P}}$$

onde:

$$\hat{V}\left(\hat{P}\right)=\sum_{h}\frac{m_h}{m_h-1}\sum_{i}\left(Z_{hi}-\frac{Z_h}{m_h}\right)^2$$

$$Z_{hi}=\frac{\sum_{j}w_{hij}\left(x_{hij}-\hat{P}_{hi}\right)}{\sum_{j}w_{hij}}$$

$$Z_h=\sum_{i}\frac{\sum_{j}w_{hij}\left(x_{hij}-\hat{P}_{hi}\right)}{\sum_{j}w_{hij}}$$

$\hat{V}\left(\hat{P}\right)$ é o estimador de variância das características de interesse da pesquisa;

$\hat{P}_{hi}$ é o estimador de proporção no i-ésimo setor do h-ésimo estrato, a saber:

$$\hat{P}_{hi}=\frac{\sum_{j}w_{hij}x_{hij}}{\sum_{j}w_{hij}}$$

$\hat{P}$ é o estimador de proporção de interesse da pesquisa.

5.4.7.4 Resultados

Os resultados da pesquisa, na forma de estimativas de proporções de domicílios, por alternativa de resposta, para cada quesito do questionário, subsidiaram a decisão do formato final de comunicação/divulgação da operação censitária para o Censo 2000.

Assim, 95,5% das pessoas investigadas informaram ter conhecimento de que havia sido realizado o Censo Experimental no município e 65,2% tinham sido elas mesmas as entrevistadas no Censo. O comercial na TV (47%), o Recenseador (29%), o comercial no rádio (22%) e o jornal (17%) foram os veículos de comunicação mais apontados pelas pessoas, através dos quais tomaram conhecimento do Censo Experimental, no município, antes ou durante a sua realização.

O comercial na TV e o próprio Recenseador, cada um com 30%, foram as formas como as pessoas compreenderam que o Censo Experimental era um ensaio para o Censo 2000. É interessante registrar que o quesito que investigava se a propaganda teve influência sobre a decisão do morador em abrir a porta para o Recenseador do IBGE apresentou estimativas de 50% para cada uma das respostas – sim ou não.

Referências

AZEVEDO, R. V.; FREITAS, M. P. S; ALBIERI, S. *Aspectos de amostragem relativos à pesquisa de avaliação da comunicação do censo experimental em Marília*. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2000. 22p.

BERCOVICH, A. O Censo 2000 no Mercosul: uma experiência transnacional. In: MIGRAÇÕES internacionais: contribuições para políticas. Brasília: Comissão Nacional de População e Desenvolvimento, 2001. p. 275-280. Trabalho apresentado no Seminário sobre Migrações Internacionais - Contribuições Para Políticas, 2000, Brasília.

BERCOVICH, A.; PEREIRA, N. M. Unidad de empadronamiento: hogar, vivienda o familia? Ventajas y desventajas del enfoque del censo brasileño. In: AMÉRICA Latina: aspectos conceptuales de los censos del 2000. Santiago de Chile: CELADE, 1999. p. 171-182.

CAMARDELLA, C. C.; XAVIER, E. *Relatório sobre a prova piloto conjunta do Censo 2000 comum dos países do Mercosul* . Puerto Iguazú, 1998.

CAZZOLLI, L. I Prueba piloto conjunta. In: TALLER DE DISCUSION DEL CENSO DE POBLACION Y VIVIENDAS 2000 EN EL MERCOSUR, 7., 1999, Santiago de Chile. Santiago de Chile: INDEC, 1999.

_____. Censo común del Mercosur: ronda del 2000: documento síntesis. In: SEMINÁRIO SOBRE O CENSO 2000 NO MERCOSUL, 8., 1999, Rio de Janeiro. Rio de Janeiro: IBGE, 1999. 20 p.

_____. *Informe de observacion en campo*. Foz de Iguaçú, 1998.

CENSO demográfico 2000: manual do recenseador – CD 1.09. Rio de Janeiro: IBGE, 2000. 151 p.

CENSO experimental 1999: agenda do supervisor. Rio de Janeiro: IBGE, 1999. 77 p.

_____: avaliação do treinamento: instrutor – TR 1.11. Rio de Janeiro: IBGE, 1999.

_____: avaliação do treinamento: treinandos – TR 1.10. Rio de Janeiro: IBGE, 1999.

_____: caderno de exercícios – TR. 04. Rio de Janeiro: IBGE, 1999. 37 p.

_____: gabarito do teste final A – TR. 08. Rio de Janeiro: IBGE, 1999. 15 p.

_____: gabarito do teste final B – TR. 09. Rio de Janeiro: IBGE, 1999. 14 p.

_____: informações gerais: censo experimental e o trabalho do recenseador. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1999. 34 p.

_____: instruções operacionais para agentes censitários municipais e coordenadores do censo experimental. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1999. 5 p.

_____: instruções para autopreenchimento do questionário básico. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1999. 4 p.

CENSO experimental 1999: agenda do supervisor. Rio de Janeiro: IBGE, 1999. 77 p.

_____: manual de coordenação – CD 1.12. Rio de Janeiro: IBGE, 1999. 13 p.

_____: manual do agente censitário municipal – CD 1.11. Rio de Janeiro: IBGE, 1999. 9 p.

_____: manual do instrutor – TR. 01. Rio de Janeiro: IBGE, 1999. 139 p.

_____: manual do recenseador – CD. 1.09. Rio de Janeiro: IBGE, 1999. 99 p.

_____: manual do supervisor - CD. 1.10. Rio de Janeiro: IBGE, 1999. 21 p.

_____: orientações ao observador. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1999. 12 p.

_____: pesquisa de avaliação da cobertura da coleta: instruções operacionais para a entrada de dados - PA 1.12. Rio de Janeiro, IBGE, 1999. 9 p.

_____: pesquisa de avaliação da cobertura da coleta: manual de acompanhamento e empastamento - PA 1.11. Rio de Janeiro, IBGE, 1999. 8 p.

_____: pesquisa de avaliação da cobertura da coleta: manual de batimento e reconciliação – PA 1.10. Rio de Janeiro: IBGE, 1999. 30 p.

_____: pesquisa de avaliação da cobertura da coleta. manual do recenseador – PA 1.09. Rio de Janeiro: IBGE, 1999. 34 p.

_____: relatório de avaliação do treinamento do censo experimental de 1999: BONITO/PA, MARILIA/SP. Rio de Janeiro: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, 1999. 24 p.

_____: roteiro de estudo para o recenseador: auto-instrução – TR. 02. Rio de Janeiro: IBGE, 1999. 54 p.

_____: teste final A – TR. 06. Rio de Janeiro: IBGE, 1999. 27 p.

_____: teste final B – TR. 07. Rio de Janeiro: IBGE, 1999. 23 p.

_____: teste inicial – TR. 03. Rio de Janeiro: IBGE, 1999. 16 p.

COCHRAN, W. G. *Sampling techniques*. 3rd ed. New York: Wiley, c1977. 428 p.

COSTA, V. G. *Relatório de trabalho de campo:* censo experimental de Marília 1999. Rio de Janeiro: IBGE, Diretoria de Geociências, 1999a.

_____. *Relatório fotográfico*: censo experimental de Marília 1999. Rio de Janeiro: IBGE, Diretoria de Geociências, 1999b.

FURLAN, S.; PIETRO, R. C. *Projeto censo 2000 nas escolas:* relatório de consultoria: parâmetros curriculares nacionais – PCN. São Paulo: Universidade de São Paulo, Departamento de Geografia, 1999.

HANSEN, M. H.; HURWITZ, W. N.; MADOW, W. G. *Sample survey methods and theory*. New York: Wiley, 1953. v.2.

MARKS, E.S. *The role of dual system estimation in census evaluation.* Washington, D.C.: U.S. Bureau of the Census, 1973, 39p.

OLIVEIRA, L. C. de S.; FREITAS, M. P. S.; BIANCHINI, Z. M. *Pesquisa de avaliação da cobertura da coleta do censo demográfico 2000:* definição do desenho amostral. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1999.

PESSOA, D. G. C e Silva, P. L. N. *Análise de dados amostrais complexos*. São Paulo: Associação Brasileira de Estatística, c1998. 187 p. Trabalho apresentado no XIII SINAPE, Caxambu, jul. 1998.

PRIMEIRA prova piloto: anteprojeto para realização da primeira prova piloto: avaliação do processo de leitura. Rio de Janeiro: IBGE, Diretoria de Informática, 1997.

_____: anteprojeto para realização da primeira prova piloto: objetivos a serem alcançados na prova piloto pelo ponto de vista da informática. Rio de Janeiro: IBGE, Diretoria de Informática, 1997.

_____: manual do recenseador. Rio de Janeiro: IBGE, 1997.

_____: relatório sobre a primeira prova piloto do censo demográfico 2000. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1997.

_____: relatório sobre a primeira prova piloto do censo demográfico 2000: treinamento. IBGE, Diretoria de Pesquisas, 1997.

SEGUNDA prova piloto do censo demográfico 2000: avaliação da aprendizagem. Rio de Janeiro: IBGE, 1998. 19 p.

_____: caderno de exercícios. Rio de Janeiro: IBGE, 1998. 58 p.

_____: manual de preenchimento do questionário básico do censo demográfico 2000. Rio de Janeiro: IBGE, 1998. 6 p.

_____: manual de preenchimento do questionário da amostra do censo demográfico 2000. Rio de Janeiro: IBGE, 1998. 8 p.

_____: manual do instrutor. Rio de Janeiro: IBGE, 1998. 71 p.

_____: manual do recenseador. Rio de Janeiro: IBGE, 1998. 71 p.

_____: manual do supervisor. Rio de Janeiro: IBGE, 1998. 14 p.

_____: prova conjunta - Mercosul: manual do recenseador. Rio de Janeiro: IBGE, 1998. 71 p.

_____: objetivos da segunda prova piloto do censo demográfico 2000: Rio de Janeiro: IBGE, Diretoria de Pesquisas e Diretoria de Informática, 1998. 5 p.

_____: quesitos da segunda prova piloto do censo demográfico 2000. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1998. 45 p.

_____: relatório de avaliação da segunda prova piloto do censo demográfico 2000: 1ª parte. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1998. 61 p.

_____: relatório de avaliação da segunda prova piloto do censo demográfico 2000: 2ª parte. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1998. 18 p.

_____: roteiro para iniciação ao estudo do manual do recenseador. Rio de Janeiro: IBGE, 1998. 38 p.

_____: tabelas de análise: pessoas portadoras de deficiência. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1998. 32 p

SILVA, E. C. de J. M. da; BARROS, Salvina Pereira. *Projeto censo 2000 nas escolas:* relatório de testagem do material: fase experimental. 1999. 48 p.

XAVIER, E.; SOUZA, L. B. de. *Relatório sobre a prova piloto conjunta do Censo 2000 comum dos países do Mercosul*. Puerto Quijarro, 1998

Anexos

Primeira prova piloto

Questionário Básico;

Questionário da Amostra 1; e

Questionário da Amostra 2

Segunda prova piloto

Questionário Básico;

Questionário da Amostra;

Folha de Coleta;

Caderneta do Setor; e

Folha de Domicílio Coletivo

Censo Experimental 1999

CD 1.01 Questionário Básico;

CD 1.02 Questionário da Amostra;

CD 1.07 Folha de Coleta;

CD 1.03 Folha de Domicílio Coletivo; e

CD 1.06 Caderneta do Setor

Pesquisa de Avaliação da Cobertura da Coleta do Censo Experimental 1999

PA 1.01 Questionário da Amostra;

CD 1.07 Folha de Coleta;

PA 1.06 Caderneta do Setor; e

Questionário de Avaliação da Comunicação no Censo Experimental

6 Desenho conceitual

São vários os aspectos que determinam o desenho conceitual de um censo demográfico. Buscando facilitar a descrição desses aspectos, optou-se por apresentar, inicialmente, as questões ligadas ao âmbito do Censo Demográfico 2000, discorrendo, a seguir, sobre as reflexões que determinaram a forma e o conteúdo da investigação de alguns temas. Completa-se a descrição do desenho conceitual com a apresentação exaustiva dos conceitos do censo e dos aspectos ligados às classificações utilizadas no tratamento das informações coletadas em aberto.

6.1 Âmbito da pesquisa

No que toca ao estabelecimento do âmbito da pesquisa, podem ser citados como relevantes os seguintes aspectos de cunho geral: o uso das informações, as recomendações internacionais, a experiência internacional e a consistência com a série histórica. A esses pontos somam-se outros, específicos que, no caso do Censo Demográfico 2000, foram: a crescente necessidade de informações de natureza demográfica e socioeconômica, detalhadas para o nível municipal; a otimização dos recursos disponíveis; e a obtenção de resultados de qualidade, com maior rapidez.

6.1.1 População investigada

O censo brasileiro adota o conceito de população residente ou “de direito”, ou seja, a população é enumerada no seu local de residência habitual. A outra alternativa seria o levantamento da população “de fato”, ou seja, no local em que se encontrava na

data de referência do censo. As características continentais do País e a utilização da informação censitária aconselham o uso da primeira alternativa.

A investigação das características dos domicílios e das pessoas neles residentes tiveram como data de referência a zero hora do dia 1º de agosto de 2000. Portanto, de acordo com esse critério, as pessoas nascidas após essa data deviam ser excluídas do censo, pois não faziam parte do domicílio, na data de referência. Já as pessoas, falecidas após a data, deviam ser incluídas, pois faziam parte do domicílio, na data de referência.

Foi decidido não se realizar simultaneamente ao Censo 2000, o levantamento dos que moravam, exclusivamente, na rua – moradores de rua. Os argumentos levantados nas discussões sobre esse aspecto foram:

- necessidade de ser estabelecido um questionário próprio para esse tipo de levantamento;

- necessidade de concentrar-se a coleta dos dados em apenas um fim de semana, visto a grande mobilidade dessa população; a escolha do fim de semana, deve-se à constatação de que existem aparentes "moradores de rua" que, nos fins de semana, retornam ao domicílio familiar e, neste caso, de acordo com as definições do censo, serão enumerados em seus domicílios;

- necessidade do levantamento ser realizado por uma equipe distinta daquela que faz a coleta no setor censitário, que precisaria, inclusive, receber treinamento especial para realizar as entrevistas.

- não-existência de informações que permitisse identificar, espacialmente, a concentração dessa população, o que faria que fosse necessário estabelecer um número muito grande de equipes, de modo a garantir-se a cobertura, o mais próximo possível da realidade; e

- a investigação dessa parcela da população, considerando a sua relevância, seria muito mais adequada, através de uma pesquisa específica, que tivesse seu âmbito e profundidade, bastante bem definidos.

Assim, para o Censo 2000, deviam ser recenseados todos os moradores em domicílios particulares – permanentes e improvisados – e coletivos, na data de referência. Por outro lado, por representarem território estrangeiro, as embaixadas, legações – representações diplomáticas de categoria inferior a embaixadas - ou consulados foram considerados unidades não-residenciais e, portanto, ainda que estivessem servindo de residência, seus moradores não foram objeto de pesquisa.

Desde 1960, no Brasil, tem-se recorrido à utilização de amostragem probabilística nos trabalhos censitários, sendo investigadas, cada vez em maior escala, informações através desse método. É sobretudo esse procedimento que vem permitindo a ampliação e o aprofundamento dos temas abordados, tendo em vista a obtenção de informações mais completas sobre as condições de vida da população.

Dessa forma, como vem sendo costume, durante a coleta do Censo Demográfico 2000 foram usados dois modelos de questionário, sendo, em cada domicílio, aplicado somente um dos modelos. O simplificado, denominado Questionário Básico, e outro, bem mais extenso e complexo, o Questionário da

Amostra, cuja aplicação foi feita em cada domicílio selecionado para a amostra do censo. Este último contém todas as perguntas do Questionário Básico, e mais um conjunto de quesitos sobre temas como educação, religião, deficiência, migração, fecundidade, trabalho e rendimento, entre outros.

A aplicação de um ou outro tipo de questionário dependia do tamanho populacional do município. Assim, para os municípios onde a população estimada em 1º de agosto de 2000 era de até 15 000 habitantes, a cada cinco domicílios recenseados, em quatro foi aplicado o Questionário Básico e em um o Questionário da Amostra – fração amostral de 20% dos domicílios. Por outro lado, para os municípios com população estimada acima de 15 000 habitantes, a cada dez domicílios recenseados, em nove foi aplicado o Questionário Básico e em um o Questionário da Amostra – fração amostral de 10% dos domicílios.

A aplicação dessas frações de amostragem fez com que a dimensão da amostra do censo resultasse robusta o suficiente, propiciando medidas de precisão adequadas para níveis geográficos variados, como forma de atender às demandas por informações municipais e, dependendo da dimensão populacional do município, até mesmo para áreas menores.

6.1.2 Temas investigados

Cabe lembrar que a definição dos temas investigados no censo, deu-se após ampla consulta a diversos seguimentos representativos da sociedade, e discussões internas envolvendo técnicos e analistas da Instituição, nas áreas de Estatística, Economia, Sociologia, Demografia, Educação, Saúde, etc. Ressalte-se, ainda, a presença da Comissão Consultiva do Censo Demográfico 2000, formada por especialistas de projeção nacional, que acompanhou todo o planejamento do censo, especialmente a definição do seu conteúdo temático de investigação.

6.1.2.1 Dados censitários

O conjunto de informações comuns aos dois questionários utilizados constituiu as informações básicas censitárias obtidas para 100% da população, o que se convencionou chamar Conjunto Universo. A investigação dessas informações básicas foi determinada em função dos seguintes aspectos:

- a construção de uma série de indicadores demográficos e socioeconômicos básicos, em diversos níveis geográficos, inclusive de setor censitário, para viabilizar estudos de planejamento intramunicipal;

- a garantia de grande parte da coleta em até três meses, pois não devia ser prolongada em função do afastamento da data de referência e do fato de que as pessoas poderiam mudar de residência;

- a liberação das informações para a sociedade com maior rapidez, em relação às do Questionário da Amostra;

- o subsídio ao processo de estimação das informações da amostra do censo, de tal forma que os pesos ou fatores de expansão para obtenção das estimativas satisfaçam à condição de igualar estimativas amostrais aos valores conhecidos do Conjunto Universo, para um grupo de variáveis auxiliares comuns, de cada área de interesse de estimação; e

- a utilização das informações censitárias para o planejamento das amostras das pesquisas domiciliares a serem realizadas no período intercensitário.

Em função desses requisitos, o conteúdo do Questionário Básico do Censo 2000 incluiu: dez quesitos de características do domicílio, nove quesitos para a pessoa responsável pelo domicílio ou para o morador individual em domicílio coletivo e seis quesitos para cada um dos demais membros do domicílio.

Embora a quantidade de informações obtidas para o Conjunto Universo seja aparentemente pequena, dizem respeito a aspectos básicos e outros específicos, de acordo com as necessidades do País, com alta relevância de demanda da informação, para níveis geográficos desagregados.

a) Características do domicílio

As informações básicas sobre as características dos domicílios são imprescindíveis para identificar, dentre outros aspectos, áreas de investimentos prioritários em habitação e saneamento básico, em níveis geográficos detalhados.

O número de banheiros foi considerado uma variável de alta relevância em função das análises efetuadas com resultados dos censos anteriores. Esses estudos mostraram grande correlação dessa variável com a renda domiciliar, o que confere à informação a possibilidade de contribuir para a estratificação socioeconômica, podendo ser usada no planejamento de amostras de pesquisas domiciliares.

Desse modo, foi investigado, para todos os domicílios, a sua espécie e, para os domicílios particulares permanentes, dentre outras, as seguintes características: tipo, condição de ocupação do domicílio, condição de ocupação do terreno, forma de abastecimento de água, forma de canalização da água, número de banheiros, existência de sanitário, tipo de escoadouro do banheiro ou sanitário e destino do lixo.

b) Características das pessoas

São as seguintes as justificativas, quanto ao que foi investigado para as pessoas no Questionário Básico:

– **Sabe ler e escrever**

Em função das grandes diferenças regionais das taxas de analfabetismo, o quesito "sabe ler e escrever" é perguntado a todos, pois trata-se de uma informação básica para o País que ainda conta com um considerável número de analfabetos. Vale acrescentar a importância de que as taxas de analfabetismo sejam conhecidas nos menores níveis geográficos, para implementação de políticas de combate ao problema.

– **Relação com o responsável pelo domicílio**

Para a estruturação dos arranjos domiciliares, a captação da relação de convivência existente entre cada pessoa moradora do domicílio e o responsável pelo domicílio foi feita nos moldes tradicionais, conservando-se praticamente as mesmas relações de parentesco pesquisadas nos censos anteriores, para manter a comparabilidade da série, através das seguintes relações: cônjuge ou companheiro(a), filho(a) ou enteado(a), pai, mãe ou sogro(a), neto(a) ou bisneto(a), irmão ou irmã, outro parente, agregado(a), pensionista, empregado(a) doméstico(a) e parente do(a) empregado(a) doméstico(a).

– **Rendimento e grau de instrução - só para a pessoa responsável ou para o morador individual em domicílio coletivo**

Desde que foi introduzido o Questionário da Amostra no Censo de 1960, o rendimento só era investigado nesse questionário, pois, como é sabido, essa informação não é trivial para levantamento junto às pessoas e também para a crítica de consistência posterior.

Para o Censo de 1991, houve forte demanda por parte de usuários especializados para que o IBGE incluísse, no Questionário Básico, informações que, sem grandes alterações de custos, pudessem ampliar o uso das informações censitárias. O principal deles, importante também para o próprio IBGE, seria ter alguma informação, mesmo que parcial, para o rendimento domiciliar e, conseqüentemente, para o do setor censitário, objetivando a definição de amostras para suas pesquisas domiciliares, até que os resultados da amostra do censo estivessem disponíveis. Cabe registrar que o valor da participação média do rendimento da pessoa responsável pelo domicílio, no total do rendimento domiciliar, costuma ser bastante significativa.

Também houve grande demanda para que se incluísse a informação sobre o grau de instrução da pessoa responsável. Neste caso, além de atender ao mesmo objetivo identificado para o quesito sobre rendimento, pois há forte correlação entre rendimento e grau de instrução, seria possível obter, de forma rápida, uma informação que possibilitasse a construção de alguns indicadores sociais relevantes, sem esperar, de dois a três anos, pelos resultados da amostra.

Assim, foram investigadas as seguintes características:

- **para o responsável pelo domicílio ou o morador individual em domicílio coletivo:** sexo, mês e ano de nascimento, idade em 31.07.2000, se sabe ler e escrever, curso mais elevado que freqüentou no qual concluiu pelo menos uma série, última série concluída com aprovação e rendimento bruto – do trabalho somado ao de outras fontes – do mês de julho de 2000; e

- **para os demais moradores do domicílio:** sexo, relação com o responsável pelo domicílio, mês e ano de nascimento, idade em 31.07.2000 e se sabe ler e escrever.

6.1.2.2 Dados investigados por amostragem

O Questionário da Amostra incluiu, além dos quesitos que constavam do Questionário Básico, outros quesitos sobre características do domicílio e das pessoas moradoras. Foram 23 quesitos de características do domicílio e 67 quesitos de características dos moradores do domicílio. A aplicação dos quesitos para cada pessoa dependeu da idade e do sexo, podendo acontecer saltos dentro do questionário. Por exemplo: os quesitos de nupcialidade, trabalho e rendimento só foram aplicados para pessoas com dez anos ou mais; e os de fecundidade, somente para as mulheres nessa mesma faixa etária.

a) Características do domicílio

Foram investigados os seguintes aspectos, no que se refere às características do domicílio:

- **espécie:** – para todos os domicílios, classificando-os em particular permanente, particular improvisado ou coletivo; e

- **para os domicílios particulares permanentes:** tipo, número de cômodos existentes no domicílio, número de cômodos que servem de dormitório para os moradores do domicílio, condição de ocupação do domicílio, condição de ocupação do terreno, forma de abastecimento de água utilizada no domicílio, forma de canalização da água, número de banheiros, existência de sanitário, tipo do escoadouro do banheiro ou sanitário, destino do lixo, existência de iluminação elétrica, existência dos bens duráveis: rádio, geladeira ou *freezer*, videocassete, máquina de lavar roupa, forno de microondas, linha telefônica instalada, microcomputador; e quantidade de televisores, automóveis para uso particular e aparelhos de ar-condicionado.

b) Características das pessoas

Quanto às características dos moradores, foi investigado o seguinte:

- **dados gerais:** sexo, relação com o responsável pelo domicílio, relação com o responsável pela família, número da família a que pertence, mês e ano de nascimento, idade em 31.07.2000, cor ou raça e religião ou culto;

- **deficiência:** existência de deficiência mental permanente que limite as atividades habituais, avaliação da capacidade de enxergar, avaliação da capacidade de ouvir, avaliação da capacidade de caminhar/subir escadas, existência de algumas deficiências físicas – paralisia permanente total, paralisia permanente das pernas, paralisia permanente de um dos lados do corpo, falta de uma das partes do corpo (perna, braço, mão, pé ou dedo polegar);

- **migração:** se sempre morou no município; tempo de moradia sem interrupção no município; se nasceu no município; se nasceu na Unidade da Federação; nacionalidade; ano em que fixou residência no Brasil; Unidade da Federação ou país de nascimento; tempo de residência ininterrupto na Unidade da Federação; Unidade da Federação ou país estrangeiro de residência anterior; onde residia em 31 de julho de 1995; município e Unidade da Federação ou país estrangeiro de residência em 31 de julho de 1995; município e Unidade da Federação ou país estrangeiro onde trabalha ou estuda;

- **educação:** se sabe ler e escrever; se freqüenta escola ou creche – rede particular ou pública; curso que freqüenta e série que freqüenta; curso mais elevado que freqüentou no qual concluiu pelo menos uma série; última série concluída com aprovação; se concluiu o curso que estudou e espécie de curso mais elevado concluído;

- **nupcialidade:** se vive ou viveu em companhia de cônjuge ou companheiro(a), natureza da última união e estado civil;

- **trabalho e rendimento – na semana de referência:** existência de trabalho remunerado; existência de trabalho remunerado do qual estava temporariamente afastado; existência de trabalho não-remunerado em ajuda a conta própria ou empregador – morador – ou como aprendiz ou estagiário; existência de trabalho não-remunerado em ajuda a empregado – morador – em atividades de cultivo, extração vegetal, criação de animais, caça,

pesca ou garimpo; existência de trabalho para o próprio consumo; número de trabalhos; ocupação; atividade e condição de ocupação no trabalho principal; se empregado pelo regime jurídico dos funcionários públicos ou militar; se contribuinte para instituto de previdência;

- **trabalho e rendimento – no mês de julho de 2000:** número de empregados – somente para empregadores; rendimento bruto no trabalho principal e nos demais trabalhos; horas trabalhadas habitualmente por semana no trabalho principal e nos demais trabalhos; se procurou trabalho; se aposentado de Instituto de Previdência Oficial; se possuía rendimentos provenientes de: aposentadoria, pensão, aluguel, pensão alimentícia, mesada, doação recebida de não-morador e programas oficiais de auxílio – renda mínima, bolsa-escola, seguro-desemprego, etc. – e de outras fontes; e

- **fecundidade:** número de filhos, por sexo; nascidos vivos até 31.07.2000; número de filhos, por sexo, nascidos vivos que continuam vivos em 31.07.2000; sexo do último filho nascido vivo até 31.07.2000; mês e ano de nascimento, ou idade presumida, do último filho nascido vivo até 31.07.2000; se o último filho nascido vivo estava vivo em 31.07.2000; e número de filhos, por sexo, nascidos mortos até 31.07.2000.

6.2 Conteúdo da investigação

Nesta parte do capítulo, a intenção é apresentar de maneira breve os motivos determinantes da linha de investigação do censo para alguns pontos ou temas selecionados. Alguns desses pontos foram, pela primeira vez, objeto de levantamento no censo; outros foram retirados da pesquisa; e ainda outros, apesar de novamente pesquisados, o foram através de formas distintas das anteriormente investigadas.

6.2.1 Investigação nos questionários

A operação censitária não é compatível com questionários extensos e de difícil aplicação junto aos informantes.

Por outro lado, as crescentes necessidades de informações de natureza demográfica e socioeconômica, em níveis geográficos cada vez mais detalhados, têm aumentado de forma considerável a demanda pela inclusão de quesitos nos questionários do censo.

A definição do conteúdo investigado no Censo 2000 levou em conta os seguintes aspectos:

- a finalidade do uso das informações, respeitadas as demandas prioritárias e atuais dos usuários dos censos demográficos;

- as recomendações internacionais e a experiência internacional;

- a consistência com a série histórica;

- a comparabilidade internacional, particularmente com o Censo Comum do Mercosul;

- a experiência e competência do quadro técnico da Instituição;

- as experiências externas disponíveis, a começar pelas reunidas na Comissão Consultiva do Censo Demográfico 2000 e nos principais usuários e especialistas temáticos; e

- o impacto nos custos da operação, quanto à inclusão da informação solicitada.

É importante observar que a maioria das decisões tomadas com respeito ao conteúdo do Questionário Básico impacta, do mesmo modo, a investigação do Questionário da Amostra.

6.2.1.1 Questionário Básico

O conteúdo do Questionário Básico inclui um conjunto de itens existentes na maioria dos censos de todo o mundo. Assim, além dos temas básicos anteriormente citados, apenas questões de alta relevância e que só podem ser mensuradas por uma operação censitária devem ser investigadas.

Condições de habitação

Foram retirados os quesitos: material predominante nas paredes externas do domicílio, material predominante no piso do domicílio, material predominante na cobertura do domicílio e paredes do domicílio. A análise dos dados das pesquisas da década e dos resultados das provas piloto apontaram que esses quesitos, isoladamente, não mais distinguiam a qualidade da habitação, a menos que fossem acompanhados de outros atributos. Especialistas em habitação consideraram que as características relacionadas ao entorno dos domicílios – calçamento, iluminação pública, identificação dos logradouros – classificariam melhor as moradias. Decidiu-se, então, pela inclusão desses quesitos na Folha de Coleta, o que permitiria uma correspondência das informações dos domicílios com seu entorno, além de reduzir o custo derivado de incluir tais quesitos no questionário.

Número de banheiros

O número de banheiros foi considerado uma variável relevante e, portanto, mantido em função das análises efetuadas com resultados dos censos anteriores. Esses resultados mostraram uma elevada correlação dessa variável com a renda domiciliar, constituindo uma informação com poder de estratificação socioeconômica.

Valor do aluguel

O quesito "valor do aluguel" é objeto do levantamento anual da Pesquisa Nacional por Amostra de Domicílios – PNAD – e, em função disso, não houve necessidade de se incluir essa investigação no censo. O IBGE conta com uma série histórica de informações que atendem às demandas existentes, não sendo necessário desagregar essa informação em nível municipal.

Telefonia fixa/celular

Após testes nas provas piloto, optou-se por investigar apenas a existência de linha telefônica instalada no domicílio, apesar da crescente proporção de pessoas com acesso à telefonia celular, mas sem linha telefônica fixa instalada

no domicílio. Para uma investigação mais abrangente sobre a telefonia, deveria ser considerado o acesso das pessoas ao telefone como meio de comunicação, levando em consideração até o telefone público, o que implicaria em ampliar ainda mais o número de perguntas do questionário.

Pessoa responsável pelo domicílio

Nos grandes levantamentos estatísticos, para conhecer as relações entre os moradores de um domicílio é tradição que, primeiro, se identifique o chefe ou responsável ou a pessoa de referência e, depois, as pessoas restantes, de acordo com sua relação com esta pessoa. A denominação que se dá a esta primeira pessoa identificada no questionário varia entre os países. No caso do Brasil, os censos e pesquisas domiciliares utilizaram, durante muitos anos, a denominação "chefe do domicílio" e "chefe da família". O termo "chefe do domicílio" ou "chefe da família" sempre esteve associado à autoridade e responsabilidade pelos negócios da família e, na maioria dos casos, a mais importante fonte de sustento. Os dados mostraram, ao longo dos anos, a predominância de pessoas do sexo masculino nessa escolha.

Com o passar dos anos e a crescente participação de todos os membros da família nas decisões de âmbito familiar, o crescimento do número de pessoas economicamente ativas por domicílio e o conseqüente compartilhamento no sustento da família, a palavra "chefe" passou a ser considerada inadequada e foi abandonada. No início da década de 1990, as novas pesquisas domiciliares lançadas pelo IBGE passaram a utilizar o termo "pessoa de referência" para identificar a primeira pessoa do questionário, a partir da qual seria verificada a relação entre os moradores do domicílio e observadas as estruturas familiares. No entanto, os pesquisadores estavam críticos com essa denominação, devido à dificuldade de entendimento e compreensão por parte do informante sobre o significado da palavra "referência". Na Primeira Prova Piloto do Censo 2000, foi usado o termo "pessoa de referência", mas a avaliação dos pesquisadores levou a alterá-lo, na segunda prova, para "pessoa responsável", termo que foi aceito pelos entrevistados e recomendado para o questionário definitivo.

As instruções de preenchimento dos questionários no Censo 2000 orientavam para que a pessoa responsável pelo domicílio fosse aquela reconhecida como tal pelos demais moradores, seja homem ou mulher, e era a primeira pessoa do domicílio a ser registrada no questionário.

Parceiros do mesmo sexo

Tratando-se das relações conjugais entre as pessoas do mesmo sexo, tentou-se na primeira prova piloto fazer essa investigação, criando-se uma categoria específica – parceiro – e o resultado foi pouco satisfatório, levando à sua retirada como categoria específica.

Idade

Esta é uma das variáveis mais robustas de um censo. Como o método de captura de dados utilizado foi o *scanner* e o reconhecimento ótico de caracteres, optou-se por perguntar a idade de dois modos diferentes: pelo mês e ano de nascimento e pela idade declarada em anos completos. Desse modo, o confronto entre essas duas informações permitiria aumentar a confiabilidade do processo de reconhecimento.

6.2.1.2 Questionário da Amostra

Após a Segunda Prova Piloto, através das informações colhidas nos relatórios dos Supervisores e observadores e da análise dos dados, ficou evidenciado que o Questionário da Amostra levado a campo era extenso e complexo em diversos segmentos, afetando a qualidade da coleta. Esta constatação veio ao encontro da necessidade de adequar os custos da operação censitária no orçamento.

Nas diversas reuniões realizadas posteriormente, com vistas à redução e adequação do Questionário da Amostra, cada tema foi analisado por especialistas da Diretoria de Pesquisas – DPE, conforme a pertinência das sugestões. Dada a necessidade de redução do questionário, sua consolidação foi feita baseando-se na combinação dos critérios: prioridade da utilização das informações; redução na complexidade da investigação; e agilidade no trabalho da coleta.

As variáveis que foram alteradas ou incorporadas no Questionário da Amostra no Censo Demográfico 2000, em relação aos censos anteriores e às provas piloto são as seguintes:

União conjugal com pessoa que não mora no mesmo domicílio

Na primeira prova piloto foi criado um quesito com a intenção de se identificar esses novos arranjos familiares. A partir das análises efetuadas, considerou-se que o censo não era o levantamento adequado para uma primeira investigação deste tema, havendo necessidade de uma pesquisa específica sobre o assunto, com mais perguntas, para essa caracterização.

Mortalidade

Assim como no Censo Demográfico 1991, optou-se por não pesquisar o quesito "Neste domicílio morava alguma pessoa que faleceu nos últimos 12 meses", uma vez que para atender às demandas dos estudos da mortalidade deveria ser feita uma investigação para todas as pessoas. Essa proposta não foi aceita, considerando o custo deste tipo de investigação.

Uma segunda proposta foi colocar este quesito na amostra. Porém, na primeira prova piloto, os entrevistadores notaram uma série de dificuldades na sua aplicação, assim como na obtenção de uma resposta confiável. Diante de tal situação, a demanda dos usuários também não seria atendida e optou-se pela retirada deste quesito.

Os níveis da mortalidade continuariam sendo obtidos de forma indireta, através das perguntas clássicas e dos padrões de mortalidade calculados a partir das informações do Registro Civil e do Sistema de Informações de Mortalidade do Ministério da Saúde, que estão disponíveis por idade.

Cor ou raça

A informação levantada pelo Censo Demográfico 2000 quanto à característica de cor ou raça da população brasileira seguiu a metodologia da auto-identificação ou autoclassificação do entrevistado, já utilizada em censos anteriores, onde se procurou captar a dinâmica das relações raciais na sociedade brasileira, utilizando-se cinco categorias, a saber: branca, preta, amarela, parda e indígena.

Para o Censo Demográfico 2000, o IBGE promoveu reuniões com usuários internos e externos dessas informações. Na oportunidade, as discussões sobre a incorporação de outras terminologias, que substituíssem as categorias

anteriormente utilizadas, foram objeto de alguns testes. As sugestões dos diversos usuários foram incorporadas nas duas provas piloto do Censo Demográfico 2000 e na Pesquisa Mensal de Emprego – PME, de 1998.

Na primeira prova piloto aplicou-se dois tipos de questionário. O primeiro investigava a origem étnico/racial[1], alertando para que a origem étnica ou racial não fosse confundida com o local do nascimento e com a pergunta "cor ou raça", segundo as seguintes categorias: branca, preta, amarela, indígena e parda. No segundo questionário, perguntava-se primeiramente o quesito da cor ou raça, com outras categorias: branca, negra, asiática, indígena e mestiça, e o mesmo quesito da origem. Os resultados revelaram que a diferença significativa ocorreu entre as denominações parda e mestiça, com um deslocamento de uma parcela de pessoas, da segunda para a primeira categoria. A ordem do quesito "cor ou raça" modificada nos questionários não influiu nas respostas, observando-se uma absoluta concentração de declaração de "origem brasileira" nos dois modelos.

Na segunda prova piloto, foi incluído o quesito "cor ou raça" com as categorias tradicionais branca, preta, amarela, parda e indígena, e a origem foi investigada através de um quesito aberto com opção de até duas declarações, caso o informante julgasse necessário. A origem estava relacionada à etnia, ou seja, à descendência de povo ou raça, e o que se pretendia era uma resposta de acordo com a identidade cultural, racial ou étnica de cada pessoa sob seu próprio ponto de vista; ao final, os resultados revelaram a grande maioria das respostas quanto à naturalidade das pessoas.

Dando prosseguimento às análises da cor ou raça para subsidiar a tomada de decisões a respeito do assunto, o IBGE, em 1998, introduziu, na PME, um conjunto de perguntas. Esta pesquisa foi aplicada nos 4 000 domicílios da PME. A investigação se deu mediante a introdução de dois quesitos, um com a investigação aberta e o outro, fechada, com a classificação utilizada pelo IBGE tradicionalmente. No primeiro foram observadas mais de 200 respostas diferentes, que apresentavam alguma semelhança com as respostas da PNAD de 1976. Os resultados confirmaram que o Brasil não tinha linhas de demarcação nítidas entre populações em termos de características étnicas, lingüísticas, culturais ou históricas, o que fazia com que qualquer tentativa de classificar as pessoas de acordo com estas categorias estivesse sujeita a grande imprecisão (SCHWARTZMAN, 1999).

Nesta pesquisa, também foi testada a investigação da origem de duas formas: um quesito aberto e outro pré-codificado, com 12 categorias. Os dados mostraram que muitas pessoas entenderam "origem" em termos raciais, e outras em termos de regiões, estados e cidades de origem, mas a maioria entendeu a pergunta em termos de nacionalidade.

A Comissão Consultiva do Censo Demográfico 2000 se reuniu no IBGE em dezembro de 1998, quando foi informada dos primeiros resultados da PME. Depois de amplo debate, os membros da comissão resolveram, por maioria, recomendar ao IBGE que mantivesse para o Censo a pergunta sobre "cor ou raça" tal como ela tem sido aplicada até esse momento, e não incluísse uma nova questão sobre origem.

[1] Pergunta - Qual(is) a(s) palavra(s) que melhor define(m) a sua origem? (assinalar todos os itens correspondentes à resposta da pessoa) - Brasileira, Africana, Portuguesa, Italiana, Japonesa, Judaica, Árabe, Indígena, Latino-americana (não considerar nesta categoria as pessoas com origem brasileira) e Outra (especifique).

Religião

As estatísticas sobre a religiosidade da população do Brasil têm nos censos brasileiros uma referência. Muitas pesquisas sobre religião são realizadas, mas o censo é a única que cobre todo território nacional e é feita regularmente desde 1940. O IBGE vem trabalhando em parceria com o Instituto Superior de Estudos de Religião – ISER – na classificação das religiões desde 1991, objetivando o conhecimento de novas religiões e a manutenção da comparabilidade com os censos anteriores. Detalhes sobre a nova classificação utilizada podem ser encontrados no item 6.4.

Pessoas portadoras de deficiência

Em observância à exigência legal exposta no artigo 17, da Lei Federal nº 7.853, de 24 de outubro de 1989, que criou a obrigatoriedade de incluir nos censos demográficos questões específicas sobre pessoas portadoras de deficiência, com o objetivo de quantificar esse grupo e conhecer sua realidade social e econômica, o IBGE, através do Censo Demográfico 1991, deu início ao estudo deste tema. Nessa ocasião havia uma única pergunta no questionário que estabelecia o número de portadores de deficiência, segmentada em mental, sensorial – cegueira e surdez – e física – paralisia total, paralisia de um dos lados, paralisia das pernas, falta de membro ou parte dele. A formulação da pergunta e o resultado obtido não satisfizeram a maioria dos usuários dessa informação.

Para o Censo 2000, o IBGE esteve em contato permanente com a Coordenadoria Nacional para a Integração da Pessoa Portadora de Deficiência – CORDE – do Ministério da Justiça, com a finalidade de promover uma reformulação das questões sobre o tema.

Nesta parceria, as questões formuladas privilegiaram o estudo das incapacidades como ponto de partida para a identificação das deficiências e do grau de comprometimento das condições físicas ou mentais das pessoas pesquisadas. Com isso, foi criada a possibilidade de se observar parte da percepção que tais pessoas têm em relação às alterações provocadas pela deficiência nas suas capacidades de realização, no seu comportamento e na sua participação social. Esta ferramenta de estudo cria condições para definição de medidas específicas, mais adequadas à realidade nacional.

Para chegar à formulação utilizada atualmente, foi realizado um teste na primeira prova piloto com diferentes elencos de perguntas, objetivando escolher aquelas que melhor captavam a população em estudo.

Na primeira prova piloto, no primeiro modelo de questionário, o quesito que investigava deficiência era semelhante ao apresentado no Censo de 1991, incorporando um quesito específico para deficiência mental.

No segundo modelo de questionário foi criada uma pergunta de filtro: "Tem algum problema físico ou mental que limite suas atividades de rotina?". Nesta pergunta não se colocou salto, ou seja, todas as pessoas continuariam a responder as perguntas específicas. Se o resultado mostrasse que o uso da pergunta-filtro era suficiente, poder-se-ia evitar a repetição das mesmas perguntas para todas as pessoas, o que facilitaria a coleta. Entretanto, os resultados mostraram que isso não ocorreu. Algumas pessoas respondiam não serem portadoras de deficiência, porém, quando questionadas quanto a sua capacidade de enxergar, de ouvir ou de caminhar/subir escadas, declaravam ter alguma ou grande dificuldade. Diante desses resultados, optou-se por eliminar esta pergunta-filtro.

Sob o consenso de que internacionalmente a definição de incapacidade não se restringe às deficiências no máximo de severidade, sendo importante caracterizar o grau da limitação, foram incorporados, no segundo questionário da primeira prova piloto, as categorias "nenhuma dificuldade", "alguma dificuldade", "grande dificuldade" e "incapaz" para todos os tipos de deficiência. Por ocasião da segunda prova piloto foi incluída uma pergunta sobre o atendimento escolar especial. Posteriormente, essa pergunta foi retirada por solicitação dos membros do grupo de trabalho.

Foram discutidas algumas questões acerca da formulação do quesito referente à deficiência mental. Segundo a CORDE, era necessário retirar os portadores de doença mental da pergunta sobre pessoas portadoras de deficiência mental. Dessa forma, optou-se por fazer a seguinte pergunta no Censo 2000: "Tem alguma deficiência mental permanente que limite as suas atividades habituais?" O manual de instrução do censo esclarecia quem deveria ser incluído como deficiente mental.

As análises dos resultados das referidas provas piloto, juntamente com a experiência internacional, levaram à formulação de questões aplicadas no Censo Demográfico 2000. Em três perguntas utilizou-se o conceito de limitação de atividades para identificar as incapacidades. As outras duas perguntas referiam-se às estruturas do corpo – diferentes tipos de paralisia e falta de perna, braço, mão, pé ou dedo polegar – ou às suas funções – retardo mental. O conceito utilizado incluindo diversos níveis de limitação de atividades permite distinguir os graus de incapacidade de enxergar, ouvir e locomover-se, além de caracterizar a população portadora de deficiência e seus aspectos socioeconômicos.

Esta conceituação é compatível com as recomendações internacionais, especialmente com a *International Classification of Functioning, Disability and Health* – ICF (2001) – divulgada pela Organização Mundial da Saúde. As equipes do IBGE e da CORDE tiveram disponível uma primeira versão desta publicação na ocasião da discussão do questionário do censo. Posteriormente, Mbogoni e Synneborn (2003) formalizaram a recomendação de utilizar a classificação internacional como marco teórico e seus conceitos e terminologias para fornecer a comparabilidade internacional das estatísticas produzidas, sempre que atendam às necessidades de informação de cada país.

Migração

O estudo do tema "Migração" sofreu algumas mudanças entre os Censos Demográficos 1991 e 2000. No primeiro, era possível quantificar e identificar movimentos migratórios entre municípios de uma mesma Unidade da Federação e municípios de Unidades da Federação diferentes, bem como a mudança na situação do domicílio.

No Censo 2000, para o indivíduo que respondeu que não morava no município onde foi recenseado desde que nasceu, pergunta-se "Há quanto tempo mora sem interrupção neste município" e se "nasceu neste município". A combinação destes quesitos permite detectar o migrante em nível municipal. Se o tempo que morou sem interrupção no município é menor que a idade declarada e respondeu que nasceu neste município, é possível ter um indicativo de migração de retorno. Este procedimento permite quantificar os movimentos migratórios, não permitindo, porém, identificar o município de procedência. A pessoa que respondeu que nasceu no município em que foi recenseada passa a responder a

pergunta em que declara o tempo ininterrupto em que morou naquela Unidade da Federação. Se esse tempo fosse inferior a dez anos, indicaria a Unidade da Federação de residência anterior. Para o indivíduo que respondeu que não nasceu no município de residência, seria investigada sua nacionalidade. Se brasileiro nato, seria questionado sobre a Unidade da Federação de nascimento, para, em seguida, responder o tempo ininterrupto na Unidade da Federação em que foi entrevistado e, se menor que dez anos, declarar a Unidade da Federação de residência anterior. Foi neste quesito que ocorreu a principal mudança da forma de se investigar o movimento migratório de última etapa, entre as duas pesquisas. Em 2000, não foi investigado o município de residência anterior, somente a Unidade da Federação ou país estrangeiro de residência anterior.

Foi retirado o quesito "lugar de residência dos filhos sobreviventes" no questionário definitivo da amostra do Censo Demográfico 2000. Os censos demográficos permitem estimar o número de imigrantes internacionais, mas existem diversas limitações para se estimar a emigração internacional. Com o propósito de contornar o problema, o Grupo de Trabalho sobre Metodologia para o Estudo da Migração Internacional, da International Union for the Scientific Study of Population – IUSSP –, explorou as possibilidades de utilização dos censos ou pesquisas dos próprios países para mensurar o número de emigrantes, bem como as características demográficas como sexo e idade. Surgiu, então, a proposta metodológica de Somoza (1977), que utiliza a informação sobre o lugar de residência dos filhos sobreviventes declarados pelas mulheres.

Somoza, em sua formulação, sugeriu a combinação de duas perguntas incluídas nos questionários dos censos ou pesquisas. A primeira refere-se ao lugar de residência das mães - no próprio país ou em outro – que pode ser uma complementação da pergunta sobre orfandade materna, e a segunda pergunta seria respondida pelas mães, ao declarar o lugar de residência dos filhos sobreviventes – no país ou no exterior.

Na segunda prova piloto, muito embora a amostra não fosse representativa para nenhum dos estados selecionados, foi introduzida, pela primeira vez no Brasil, a pergunta sobre o número de filhos sobreviventes que residiam no exterior. O objetivo da incorporação desse quesito era evidentemente observar as freqüências das respostas e avaliar a pertinência de mantê-lo no questionário definitivo do Censo 2000, com o propósito de se obter um valor mínimo para a estimativa dos emigrantes internacionais. Mesmo sem crítica alguma, e levando-se em consideração as limitações já apontadas, um exame dos dados tabulados revelou que o potencial de aproveitamento dos resultados podia ser elevado.

Entretanto, devido à necessidade de uma racionalização do Questionário da Amostra do Censo 2000 e ao fato de a experiência dos países que incluíram a pergunta em seus censos ser pouco conhecida, como apontado por Chackiel (1998), foi consensual que não era o momento oportuno para se investigar este quesito através de um censo demográfico. Como a pergunta sobre o lugar de residência dos filhos sobreviventes está inserida no bloco que permite estimar a fecundidade e a mortalidade infanto-juvenil, em caso de uma eventual presença de elevado número de não-respostas, todo o bloco poderia ser prejudicado. Além disso, haveria a necessidade de se incluir mais dois quesitos a serem aplicados a toda a população: um referente à condição de orfandade e outro correspondente ao lugar de residência das mães vivas, no País ou no exterior.

Com isso, foi sugerido que esta forma de mensurar parte da emigração internacional fosse testada em alguma das PNADs, na primeira década do século XXI, uma vez que a pesquisa tem representatividade nacional.

Deslocamento para trabalho ou estudo

A partir dos resultados do Censo Experimental, decidiu-se investigar o deslocamento para trabalho ou estudo num único quesito, incorporado no bloco de migração: "município e Unidade da Federação ou país estrangeiro onde trabalha ou estuda". Foram estabelecidos critérios de prioridade em caso da possibilidade de mais de uma resposta.

A decisão sobre a agregação foi tomada pela necessidade de reduzir o questionário e pela possibilidade de chegar-se ao objetivo, a partir das outras perguntas do questionário. A informação do deslocamento pendular é utilizada para as definições de áreas metropolitanas e aglomerações urbanas em todo o mundo.

Trabalho e Rendimento

O conceito de trabalho é muito amplo e sua captação implica um grande número de quesitos. Após ampla discussão com especialistas da área, optou-se por restringir a investigação ao trabalho na semana de referência, a fim de garantir comparabilidade internacional.

Com a finalidade de melhorar a captação do número de pessoas ocupadas, no Censo 2000 a pergunta que investigou trabalho foi desmembrada em cinco quesitos: o primeiro investigou o trabalho remunerado na semana de referência; o segundo tratou das pessoas que estavam temporariamente afastadas do trabalho; o outro quesito investigou o trabalho sem remuneração, tanto para aprendiz ou estagiário; o quarto quesito pesquisou o trabalho sem remuneração na atividade de cultivo, extração vegetal, criação de animais, caça, pesca ou garimpo – este desmembramento buscou explicitar de forma clara as possíveis atividades relacionadas a esse tipo de trabalho; e, por último, foi investigado o trabalho destinado à alimentação de pessoas moradoras no domicílio – trabalho para o próprio consumo.

Ao ser colocado no questionário o quesito que investiga o "trabalho para o próprio consumo", excluiu-se o que pesquisava o trabalho na construção para o próprio uso – forma de captação de trabalho adotada na segunda prova piloto – devido à baixa participação deste tipo de trabalho na população ocupada, 0,3% conforme demonstram os resultados das PNADs, de 1992 a 1997.

Alguns outros quesitos também foram testados e retirados, a saber: emprego temporário, recebimento do vale-transporte e auxílio-refeição, município onde se localizava o trabalho que exercia, e existência de algum trabalho um ano antes da data de referência.

A identificação de "emprego temporário" foi retirada porque este tipo de investigação demanda um aprofundamento da pesquisa, que implica um aumento do número de perguntas, mais compatíveis, portanto, com outro tipo de pesquisa.

A investigação, para empregadores, sobre o número de empregados ocupados, foi incluída, tendo em vista a sua utilidade para o dimensionamento da amostra da pesquisa sobre a economia informal. Reestruturou-se também o quesito "posição na ocupação", desagregando-se segundo a existência de carteira de trabalho assinada e o trabalhador doméstico.

No caso dos rendimentos, resolveu-se incluir no levantamento aqueles provenientes de programas oficiais de auxílio, ou seja, pagos por órgão governamental, os quais passaram a fazer parte da pergunta. Nesse mesmo bloco, decidiu-se agregar as categorias referentes à aposentadoria e à pensão, porque a maior parcela corresponde à aposentadoria e a fonte pagadora é a mesma. Da mesma forma, juros e dividendos de aplicações financeiras foram incluídos na categoria outros rendimentos.

O quesito "recebimento de vale-transporte" não foi incluído, uma vez que se trata de uma investigação que já é feita com maior profundidade na PNAD.

Fecundidade

As principais alterações verificadas na forma de investigação da fecundidade, nos últimos censos demográficos brasileiros, deram-se, basicamente, no quesito que permite captar a fecundidade atual das mulheres. Os filhos tidos nascidos vivos, filhos tidos nascidos mortos e os filhos tidos nascidos vivos sobreviventes, todos investigados na data dos recenseamentos, vêm sendo pesquisados regularmente desde 1970, sendo que o sexo dos filhos somente passou a ser investigado a partir de 1980.

Com o surgimento da técnica desenvolvida por Brass e outros (1968) para estimar o nível da fecundidade feminina, aos questionários dos censos demográficos foi incorporada a pergunta feita ao conjunto de mulheres, a partir de certa idade, que permitia quantificar a intensidade e a estrutura por idade da fecundidade em um período de tempo próximo da data do censo. Foi assim, através de formulações distintas, em 1970, 1980 e 1991.

No questionário do Censo Demográfico 2000, a fecundidade das mulheres de 10 anos ou mais foi investigada de maneira semelhante à que foi realizada em 1991. As principais diferenças consistem:

- na introdução do quesito que indaga separadamente o sexo do último filho tido nascido vivo;

- na opção de declarar a idade presumida do último filho tido nascido vivo, caso a(o) informante desconheça o mês e ano de nascimento; e

- investigação da sobrevivência deste último filho tido nascido vivo, quesito que foi adicionado ao questionário com o intuito de se obter uma medida direta da mortalidade infantil.

Educação

Para aqueles que se declaravam estudantes, foi investigada a rede de ensino com a finalidade de distinguir entre a freqüência às escolas públicas e particulares, visando verificar o atendimento deste serviço aos diversos grupos socioeconômicos.

Por outro lado, com a inclusão da investigação de freqüência à creche, foi eliminado o limite inferior de idade que prevalecia nos censos anteriores.

O quesito "condição de conclusão" do curso foi incluído a partir da primeira prova piloto, a fim de facilitar os trabalhos de crítica e tabulação das informações, que eram feitos com o auxílio do quesito espécie do curso concluído, que foi mantido na investigação.

6.2.2 Investigação da Folha de Coleta e na Caderneta do Setor

Para o Censo Demográfico 2000 foram elaborados e incluídos na Caderneta do Setor e na Folha de Coleta[2] quesitos referentes à caracterização do entorno dos domicílios, que estão associados à existência de Iluminação Pública, Identificação e Pavimentação/Calçamento nos logradouros. Esses quesitos foram elaborados com o objetivo de atender à demanda de usuários, entre eles tendo destaque os órgãos de planejamento urbano, já que tais perguntas constituem informações de grande importância para o estudo e para a definição de políticas públicas.

Por outro lado, a elaboração dos quesitos teve como objetivo constituir o elenco de indicadores que seriam utilizados para identificar as áreas com características de exclusão da cidade formal[3], como os aglomerados subnormais - favelas e áreas assemelhadas.

O procedimento metodológico para definição dos quesitos correspondeu aos seguintes passos:

- levantamento bibliográfico realizado pelos técnicos da Diretoria de Geociências – DGC e Diretoria de Pesquisas – DPE; e

- reuniões entre técnicos das dessas diretorias, para avaliação das características fundamentais, levantadas na literatura, que poderiam compor os quesitos para caracterização do entorno dos domicílios. Foram considerados aqueles aspectos associados à falta de acesso aos serviços públicos essenciais, como infra-estrutura urbana. Desta forma, aspectos como o traçado tortuoso e a limpeza das ruas, envolvendo varredura, também foram arrolados em um primeiro momento.

A partir da avaliação e discussão realizada no IBGE, foram mantidas aquelas características referentes à Pavimentação, Iluminação Pública e Identificação dos Logradouros. A seleção final destes três quesitos foi realizada considerando a viabilidade de sua operacionalização, ou seja, deveriam ser passíveis de coleta em campo através da observação direta do Recenseador.

Outro momento envolveu a discussão sobre a escolha do instrumento de coleta, onde seriam pesquisados os três quesitos. Como forma de encaminhamento desta questão foi sugerida pela DPE, a elaboração de uma proposta que considerasse a inclusão desses quesitos nos questionários, na Folha de Coleta e na Caderneta do Setor. Após amplas discussões, nas diversas esferas de decisão, optou-se por manter os quesitos nos dois últimos instrumentos. A questão dos recursos, da operacionalização em campo e da viabilidade da tabulação final dos resultados foram aspectos que influenciaram na decisão adotada, quanto à escolha dos quesitos e dos instrumentos de coleta.

Para o ajuste final da forma de apresentação dos quesitos, assim como para a definição das orientações que seriam incluídas no Manual do Recenseador, foram realizadas reuniões com a DPE, com o objetivo de adequar os conceitos propostos à sua operacionalização, na fase da coleta das informações.

---

[2] Constituem instrumentos de coleta que são preenchidos pelo recenseador. Ver 9.3.

[3] Projeto caracterização dos setores de baixa renda (Diretoria de Geociências/Departamento de Geografia).

A etapa seguinte consistiu no teste dos quesitos no Censo Experimental realizado no município de Marília, São Paulo. Através desse procedimento, buscou-se avaliar o trabalho de coleta da informação de caracterização do entorno dos domicílios (ver 5.4.4.4).

A organização dessas perguntas nos dois instrumentos de coleta e a orientação para a apuração das informações em campo possuem algumas diferenças.

Para caracterizar o entorno dos domicílios na Caderneta do Setor, o Recenseador deveria, ao final de sua coleta, indicar a existência total, existência parcial ou inexistência de cada um dos itens investigados, marcando uma das três alternativas categorizadas, conforme orientação do manual; e

No caso da Folha de Coleta, apenas no quesito referente ao calçamento/pavimentação houve a alternativa de existência parcial. Nesse instrumento, o Recenseador deveria fazer a Caracterização do Logradouro, após percorrer a face ou o trecho do logradouro.

6.3 Conceitos

Os conceitos e formas de investigação utilizados no censo foram definidos levando em conta, prioritariamente, as necessidades dos usuários, especialmente das instâncias dos governos locais e central, assim como os órgãos de pesquisa não-governamentais. Foram levadas também em consideração as recomendações internacionais (PRINCIPLES..., 1998) e a comparabilidade com as informações fornecidas pelos demais países. Neste sentido, o trabalho ampliado em conjunto com os países do Mercosul levou a uma convergência, sempre que possível, dos conceitos das variáveis, de forma a permitir que as tabulações comuns fossem realmente comparáveis.

São a seguir apresentados todos os conceitos, definições e formas de investigação utilizados no censo, compreendendo as etapas de elaboração da Base Operacional e Geográfica, coleta dos dados e divulgação das informações.

6.3.1 Base operacional e Geográfica

Seguem os conceitos, correspondentes à divisão territorial e a outras áreas de divulgação e apuração e à delimitação das áreas de coleta.

6.3.1.1 Divisão territorial brasileira

Grandes Regiões

São formadas pela agregação de unidades federadas. Há cinco Grandes Regiões no Brasil: Norte, Nordeste, Sudeste, Sul e Centro-Oeste. A divisão em grandes regiões foi instituída em 1970 e sofreu alterações nas décadas de 1970 e 1980, quanto à extinção, criação e alteração de categoria de algumas unidades federadas. As cinco grandes regiões brasileiras assim estão compostas atualmente:

- Norte – Rondônia, Acre, Amazonas, Roraima, Pará, Amapá e Tocantins;

- Nordeste – Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe e Bahia;

- Sudeste – Minas Gerais, Espírito Santo, Rio de Janeiro e São Paulo;

- Sul – Paraná, Santa Catarina e Rio Grande do Sul; e

- Centro-Oeste – Mato Grosso do Sul, Mato Grosso, Goiás e Distrito Federal.

Unidades da Federação - Distrito Federal e Estados

Distrito Federal

É a unidade autônoma onde tem sede o Governo Federal com seus poderes Executivo, Legislativo e Judiciário. Tem as mesmas competências legislativas reservadas aos estados e municípios e é regido por lei orgânica, sendo vedada sua divisão em municípios.

Brasília é a Capital Federal.

Estados

São as unidades de maior hierarquia dentro da organização político-administrativa do País; são subdivididos em municípios e podem incorporar-se entre si, subdividir-se ou desmembrar-se para se anexarem a outros, ou formarem novos estados ou territórios federais, mediante aprovação da população diretamente interessada, através de plebiscito, e do Congresso Nacional, por lei complementar. Organizam-se e regem-se pelas constituições e leis que adotarem, observados os princípios da Constituição Federal.

A localidade que abriga a sede do governo denomina-se Capital.

Municípios

Constituem as unidades autônomas de menor hierarquia dentro da organização político-administrativa do Brasil. Sua criação, incorporação, fusão ou desmembramento se faz por lei estadual, observada a continuidade territorial, a unidade histórico-cultural do ambiente urbano e os requisitos previstos em lei complementar estadual. Estas transformações dependem de consulta prévia às populações diretamente interessadas, através de plebiscito. Regem-se por leis orgânicas, observados os princípios estabelecidos na Constituição Federal e na Constituição do Estado onde se situam, e podem criar, organizar e suprimir distritos, observada a legislação estadual.

Cidade

É a localidade onde está sediada a Prefeitura Municipal. É constituída pela área urbana do distrito-sede e delimitada pelo perímetro urbano estabelecido por lei municipal.

Distritos

São unidades administrativas dos municípios. Sua criação, desmembramento ou fusão se faz por lei municipal, observada a continuidade territorial e os requisitos previstos em lei complementar estadual. Podem, a depender da legislação estadual, ser subdivididos, conforme o caso, em subdistritos, regiões administrativas, zonas e similares.

Vila

É a sede do distrito que não abriga a cidade. É delimitada pelo perímetro urbano definido, por lei municipal, como a área urbana do distrito que não seja sede do município.

Regiões Administrativas, Subdistritos e Zonas

São unidades administrativas municipais, normalmente estabelecidas nas grandes cidades, criadas através de leis ordinárias das Câmaras Municipais e sancionadas pelo prefeito.

Bairros e Similares

São subdivisões intra-urbanas legalmente estabelecidas através de leis ordinárias das Câmaras Municipais e sancionadas pelo prefeito.

Regiões Metropolitanas

São conjuntos de municípios de uma mesma Unidade da Federação, reunidos, por legislação estadual, segundo critérios estabelecidos pelo próprio estado. A Constituição Federal de 1988 outorgou aos estados o direito de alterar as Regiões Metropolitanas então existentes e de criar outras. Até a data de referência do Censo 2000 – 1º de agosto de 2000 – havia 22 Regiões Metropolitanas em 15 Unidades da Federação: Belém, Grande São Luís, Fortaleza, Natal, Recife, Maceió, Salvador, Belo Horizonte, Vale do Aço, Grande Vitória, Rio de Janeiro, São Paulo, Baixada Santista, Campinas, Curitiba, Londrina, Maringá, Vale do Itajaí, Norte/Nordeste Catarinense, Florianópolis, Porto Alegre, Goiânia.

Regiões Integradas de Desenvolvimento

São conjuntos de municípios cuja criação baseia-se no princípio de cooperação entre os diferentes níveis de governo — federal, estadual e municipal. Podem ser compostas por municípios de unidades federadas diferentes. No Diário Oficial da União, de 20 de fevereiro de 1998, foi publicada a Lei complementar nº 94, de 19 de fevereiro de 1998, autorizando a criação da Região Integrada de Desenvolvimento do Distrito Federal e Entorno - RIDE.

Mesorregiões Geográficas

As Mesorregiões Geográficas são formadas por conjuntos de municípios contíguos, pertencentes à mesma unidade da federação, que apresentam uma identidade regional originada a partir de formas de organização do espaço geográfico definidas pelas dimensões socioeconômica, natural e histórica, assim como pela rede de comunicação e de lugares que configuram uma articulação espacial. Foram instituídas pela Resolução do Presidente do IBGE nº 11, de 5 de junho de 1990.

Microrregiões Geográficas

As Microrregiões Geográficas são conjuntos de municípios contíguos, definidas como partes das mesorregiões que apresentam especificidades quanto à organização do espaço. Sua delimitação leva em conta, além das dimensões formadoras das mesorregiões, a vida de relações em nível local, pela possibilidade de atendimento às suas populações, por parte dos setores sociais básicos e do comércio varejista e atacadista. Foram instituídas pela Resolução do Presidente do IBGE nº 11, de 5 de junho de 1990.

6.3.1.2 Outras Áreas de Divulgação e de Apuração

Área urbana

Área interna ao perímetro urbano de uma cidade ou vila, definida por lei municipal. Para as cidades ou vilas onde não existe legislação que regulamente essas áreas, deve-se estabelecer um perímetro urbano para fins da coleta censitária, cujos limites devem ser aprovados pelo prefeito local.

Área urbana isolada

Área definida por lei municipal e separada da sede municipal ou distrital por área rural ou por um outro limite legal.

Área rural

Área externa ao perímetro urbano.

Área urbanizada de cidade ou vila

São classificadas como área urbanizada aquelas legalmente definidas como urbana caracterizada por construções, arruamentos e intensa ocupação humana; as áreas afetadas por transformações decorrentes do desenvolvimento urbano; e aquelas reservadas à expansão urbana.

Área de ponderação

É uma unidade geográfica, formada por um agrupamento mutuamente exclusivo de setores censitários, para aplicação dos procedimentos de calibração das estimativas das informações pesquisadas no Questionário da Amostra com as investigadas para a população como um todo.

Área não-urbanizada de cidade ou vila

São classificadas como área não-urbanizada aquelas legalmente definidas como urbanas, caracterizadas por ocupação predominantemente de caráter rural.

Área especial

É a área legalmente definida subordinada a um órgão público ou privado, responsável pela sua manutenção, onde se objetiva a conservação ou preservação da fauna, da flora e de monumentos culturais, a preservação do meio ambiente e das comunidades indígenas.

Os principais tipos de áreas especiais são:

Parques - Nacional, Estadual e Municipal

São áreas de proteção e preservação de unidades importantes, ou sistemas completos de valores naturais ou culturais e de proteção de recursos genéticos, onde se desenvolve a educação ambiental, oferecem-se oportunidades para a recreação pública e se proporcionam facilidades para a investigação e outros fins de índole científica.

Reservas ecológicas

São áreas de preservação permanente, mencionadas no artigo 18 da Lei nº 6.938, de 31 de agosto de 1981. Compreende as florestas e as demais formas de vegetação natural de preservação permanente – Decreto nº 89.336, de 31 de janeiro de 1984.

Estações ecológicas

São áreas representativas de ecossistemas brasileiros, destinadas à realização de pesquisas básicas e aplicadas à ecologia, à proteção do ambiente natural e ao desenvolvimento da educação conservacionista – Lei nº 6.902, de 27 de abril de 1981.

Reservas florestais ou Reservas de recursos

São áreas não habitadas, de difícil acesso e em estado natural, dos quais ainda se carece de conhecimento e tecnologia para o uso racional dos recursos ou quando as prioridades nacionais em matéria de recursos humanos e financeiros impedem investigações de campo, avaliação e o desenvolvimento delas no momento.

Reservas biológicas

São áreas que visam a proteger a natureza - espécies ou comunidades - e a manter o processo em um estado sem perturbações, visando à proteção de amostras ecológicas representativas do meio ambiente natural para estudos científicos, monitoramento ambiental, educação científica, e para manter recursos genéticos em um dinâmico e evolucionário estágio.

Áreas de relevante interesse ecológico

São áreas que possuem características naturais extraordinárias, ou abrigam exemplares raros da biota regional, exigindo cuidados especiais de proteção por parte do Poder Público – Decreto nº 89.336, de 31 de janeiro de 1984.

Áreas de proteção ambiental

São áreas de interesse para a proteção ambiental, a fim de assegurar o bem-estar das populações humanas e conservar ou melhorar as condições ecológicas locais.

Áreas de preservação permanente

São áreas de pouca extensão, geralmente inferior a 500 ha, com uma cobertura natural a ser preservada. Corresponde, na maioria das vezes, a uma pequena reserva florestal para preservação de recursos hídricos.

Monumentos naturais

São áreas que contêm uma ou mais características naturais importantes, de relevante significado nacional, como formações geológicas, lugares naturais únicos, espécies da flora e fauna, ou hábitat que, pela sua raridade, possam estar ameaçados ou devam ser protegidos.

Monumentos culturais

São áreas que possuem alguma característica arqueológica, histórica ou cultural, de interesse nacional ou internacional. Os sítios históricos tombados e os sambaquis, entre outros, estão neste tipo. Podem constituir áreas estatais ou terras de propriedade privada, onde se especifiquem claramente as práticas de uso da terra.

Áreas indígenas

São áreas ocupadas ou habitadas por índios não-aculturados, ou em incipiente processo de aculturação – Decreto nº 94.946, de 23 de setembro de 1987.

Colônias indígenas

São áreas ocupadas ou habitadas por índios aculturados, ou em adiantado processo de aculturação – Decreto nº 94.946, de 23 de setembro de 1987.

Reservas indígenas

São áreas destinadas a servir de hábitat a grupo indígena com os meios suficientes a sua subsistência – Decreto nº 84.337, de 21 de dezembro de 1979.

Parques indígenas

São áreas contidas em terra de posse de índios, cujo grau de integração permite assistência econômica, educacional e sanitária dos órgãos da União em que se preservaram a flora e a fauna e as belezas naturais da região – Decreto nº 82.263, de 13 de setembro de 1978.

Terras indígenas

São áreas ainda em fase de regularização de posse da terra e, portanto, sem denominação definitiva quanto ao tipo de área.

Aglomerado rural

É um agrupamento de população considerado a partir de um conjunto de edificações adjacentes e com características de permanência, situado em área legalmente definida como rural, que forma:

- uma área continuamente construída com arruamentos reconhecíveis ou disposta ao longo de uma via de comunicação;

- um agrupamento de edificações com mais de 50 unidades domiciliares, número que estaria relacionado a um montante de população superior a 250 habitantes.

Os aglomerados rurais estão classificados em dois grandes tipos:

- aglomerados rurais do tipo "extensão urbana" que possuem caráter urbano por definição; e

- aglomerados rurais "isolados", que podem ser de natureza urbana ou rural.

Aglomerados rurais do tipo "extensão urbana"

São os assentamentos situados em áreas fora do perímetro urbano legal, mas desenvolvidos a partir da expansão de uma cidade ou vila, ou por elas englobados em sua expansão. Por constituírem uma simples extensão da área efetivamente urbanizada, atribui-se, por definição, caráter urbano aos aglomerados rurais deste tipo. Tais assentamentos podem ser constituídos por loteamentos já habitados, conjuntos habitacionais, aglomerados de moradias ditas subnormais ou núcleos desenvolvidos em torno de estabelecimentos industriais, comerciais ou de serviços.

Aglomerados rurais isolados

São os assentamentos situados em área legalmente definida como rural, que atendam aos critérios de tamanho e densidade anteriormente estipulados e que se encontrem separados do perímetro urbano legal de uma cidade ou vila, ou de um aglomerado do tipo "extensão urbana" por uma distância igual ou superior a 1km.

Os aglomerados rurais "isolados" devem ser classificados em três subtipos: Povoados, Núcleos e Outros aglomerados rurais isolados.

Povoado

É o aglomerado rural isolado que corresponde a aglomerados sem caráter privado ou empresarial, ou seja, não vinculados a um único proprietário do solo (empresa agrícola, indústrias, usinas, etc.), cujos moradores exercem atividades econômicas, quer primárias (extrativismo vegetal, animal e mineral; e atividades agropecuárias), terciárias (equipamentos e serviços), ou, mesmo, secundárias (industriais em geral), no próprio aglomerado ou fora dele.

Núcleo

É o aglomerado rural isolado vinculado a um único proprietário do solo – empresa agrícola, indústria, usina, etc. – dispondo ou não dos serviços ou equipamentos definidores dos povoados. É considerado, pois, como característica definidora deste tipo de aglomerado rural isolado seu caráter privado ou empresarial.

Outros aglomerados rurais isolados

São os aglomerados que não dispõem, no todo ou em parte, dos serviços ou equipamentos definidores dos povoados e que não estão vinculados a um único proprietário – empresa agrícola, indústria, usina, etc.

Aldeia Indígena

É um agrupamento de, no mínimo, 20 habitantes indígenas em uma ou mais moradias.

6.3.1.3 Conceitos associados à delimitação das unidades territoriais de coleta

Unidade territorial de coleta

É a unidade de controle cadastral formada por área contínua, situada em um único quadro urbano ou rural, com dimensão e número de domicílios ou de estabelecimentos que permitam levantamento das informações por um único Agente Credenciado, segundo cronograma estabelecido. O Setor Censitário é a unidade territorial de coleta do Censo Demográfico 2000.

Limite da unidade territorial de coleta

É a linha de contorno imaginária, definida por uma seqüência de acidentes topográficos naturais ou artificiais.

Descrição do perímetro da unidade territorial de coleta

É a relação de acidentes topográficos naturais ou artificiais, arrolados de forma seqüencial, que definem a linha imaginária do contorno da área do setor, garantia para que o Recenseador não invada a unidade territorial de coleta de outro Recenseador ou omita parte da coleta da área sob sua responsabilidade.

Número da unidade territorial de coleta

É a designação utilizada para identificá-la em relação a outras. Tem como objetivo permitir a referência de diversas informações por unidade territorial de coleta, sendo utilizado como chave de recuperação dos diversos cadastros da base operacional.

Mapa Municipal Digital Estatístico (MMDE)

Refere-se ao mapa de um município em escala que varia de 1:50 000 a 1: 250 000, contendo a visão geral do município; mostrando bem detalhadamente sua área rural com a identificação dos limites de distritos e subdistritos, os principais acidentes físicos, localidades, fazendas, etc.; e identificando áreas urbanas, ou seja, as cidades e as vilas, como pequenas áreas sem definição de ruas, quadras, etc., com a malha de setores rurais.

Mapa Urbano Digital Estatístico (MUDE)

Refere-se ao conjunto de plantas urbanas, em escalas que variam de 1:2 000 a 1:10 000, que compõe as áreas urbanas de um município, com a identificação de suas ruas, avenidas, e eventuais marcos de referência – prédios públicos, igrejas, clubes, etc.; limites de bairros, subdistritos, regiões administrativas, perímetro urbano, etc., e a malha de setores urbanos.

Logradouro

É o espaço livre, inalienável, destinado à circulação pública de veículos e de pedestres, e reconhecido pela municipalidade, que lhe confere denominação oficial. São as ruas, travessas, becos, avenidas, praças, pontes, etc.

Localidade

Todo lugar do território nacional onde exista um aglomerado permanente de habitantes.

Situação do setor

É o código que identifica a localização da área do setor em relação ao quadro urbano ou rural (ver quadro 7.2 *Códigos de situação e tipo do setor 2000* no Capítulo 7).

Tipo do setor

É o código que classifica o setor segundo o conceito que orientou sua delimitação (ver quadro 7.2 *Códigos de situação e tipo do setor 2000* no Capítulo 7).

Setor comum ou não-especial

Não apresenta características que tornam necessário um tratamento diferenciado. Corresponde à maior parte dos setores censitários.

Setor especial

É aquele que apresenta características que tornam necessário um tratamento diferenciado de coleta em relação aos setores comuns ou não-especiais. Foram considerados sete tipos de setores especiais: 1 – Setor Especial de Aglomerado Subnormal, 2 – Setor Especial de Quartéis, Bases Militares, 3 – Setor Especial de Alojamento, Acampamentos, 4 – Setor Especial de Embarcações,

Barcos, Navios, 5 – Setor Especial de Aldeia Indígena, 6 – Setor Especial de Penitenciárias, Colônias Penais, Presídios, Cadeias, 7 – Setor Especial de Asilos, Orfanatos, Conventos, Hospitais.

Setor Especial de Aglomerado Subnormal

É um conjunto constituído por um mínimo de 51 domicílios, ocupando ou tendo ocupado até período recente, terreno de propriedade alheia – pública ou particular – dispostos, em geral, de forma desordenada e densa, e carentes, em sua maioria, de serviços públicos essenciais.

Podem se enquadrar, observados os critérios de padrões de urbanização e/ou de precariedade de serviços públicos essenciais, nas seguintes categorias:

- invasão;

- loteamento irregular ou clandestino; e

- áreas invadidas e loteamentos irregulares ou clandestinos regularizados em período recente.

Setor especial de aldeia indígena

É um agrupamento de, no mínimo, 20 habitantes indígenas em uma ou mais moradias.

6.3.2 Períodos de referência

Os períodos de referência utilizados para a investigação das informações censitárias foram: a data de referência, a semana de referência, o período de referência de 30 dias e o mês de referência.

Data de referência

A data de referência foi a zero hora do dia 1º de agosto de 2000.

Mês de referência

O mês de referência foi julho de 2000.

Período de referência de 30 dias

O período de referência de 30 dias compreendeu de 30 de junho a 29 de julho de 2000.

Semana de referência

A semana de referência compreendeu o período de 23 a 29 de julho de 2000.

6.3.3 Entorno do domicílio

Pela primeira vez no censo, foi feita a caracterização do entorno do domicílio, através da investigação das seguintes características: identificação, iluminação pública e calçamento/pavimentação; essa caracterização foi feita, separadamente, para o logradouro e para o setor.

Logradouro

É o espaço livre, inalienável, destinado à circulação pública de veículos e de pedestres, e reconhecido pela municipalidade, que lhe confere denominação oficial. São as ruas, travessas, becos, avenidas, praças, pontes, etc.

Caracterização do logradouro

A face ou o trecho do logradouro percorrido pelo Recenseador foi classificado segundo as seguintes características: identificação, iluminação pública, e calçamento/pavimentação.

Identificação

A face ou o trecho do logradouro percorrido pelo Recenseador foi considerado como tendo identificação, quando existissem, em qualquer dos lados, sinais visíveis com nomes e/ou numeração, como placas oficiais ou outra forma de identificação.

Iluminação pública

A face ou o trecho do logradouro percorrido pelo Recenseador foi considerado como tendo iluminação pública, quando existissem, em qualquer dos lados, pelo menos um ponto fixo – poste – de iluminação pública.

Calçamento/pavimentação

A existência de calçamento/pavimentação, na face ou no trecho do logradouro percorrido pelo Recenseador, foi investigada da seguinte forma:

- Total – quando, em toda a face ou em todo o trecho do logradouro percorrido, existisse calçamento/pavimentação;
- Parcial – quando, em parte da face ou do trecho do logradouro percorrido, não existisse calçamento/pavimentação; e
- Não existe – quando, em toda a face ou em todo o trecho do logradouro percorrido, não existisse calçamento/pavimentação.

Não foi considerado como calçamento/pavimentação: barro, terra, pó-de-pedra, areia e brita.

Caracterização do setor

O setor, após o término de sua coleta, em função da existência de cada uma das características – identificação, iluminação pública e calçamento/pavimentação – foi classificado da seguinte maneira:

- Total – quando o setor possuísse, em todos os seus logradouros, a característica investigada;
- Parcial – quando apenas parte dos logradouros do setor possuísse a característica indicada; e
- Não existe – quando todos os logradouros do setor não possuíssem a característica indicada.

6.3.4 Domicílio

Os conceitos e formas de investigação, referentes à investigação das características do domicílio, são apresentados a seguir, em ordem alfabética.

Automóvel para uso particular

Considerou-se como tendo automóvel para uso particular o domicílio particular permanente em que, na data de referência, um de seus moradores possuísse automóvel de passeio ou veículo utilitário para passeio ou locomoção dos membros do domicílio para o trabalho.

Incluiu-se como de uso particular o automóvel utilizado para desempenho profissional de ocupações como: motorista de táxi, vendedor que tem necessidade de transportar amostras de mercadoria para atender ou solicitar pedidos, etc., desde que fosse utilizado, também, para passeio ou locomoção de membros da família.

Banheiro

Considerou-se como banheiro o cômodo do domicílio particular permanente que, na data de referência, dispunha de chuveiro ou banheira e aparelho sanitário, e era de uso exclusivo dos moradores do domicílio.

Canalização de água

Pesquisou-se a existência, na data de referência, de canalização de água para o domicílio particular permanente ou para o terreno ou a propriedade em que se localizava. Classificou-se esta situação como:

- Canalizada em pelo menos um cômodo – quando o domicílio era servido de água canalizada com distribuição interna para um ou mais cômodos;

- Canalizada só na propriedade ou terreno – quando a água utilizada no domicílio chegava canalizada à propriedade ou ao terreno em que se localizava o domicílio, sem haver distribuição interna em pelo menos um cômodo; ou

- Não-canalizada – quando a água utilizada no domicílio não chegava canalizada à propriedade ou ao terreno em que se localizava o domicílio e não havia distribuição interna.

Cômodo - parte de um domicílio

Dependência do domicílio particular permanente coberta por um teto e limitada por paredes, inclusive banheiro e cozinha de uso exclusivo dos moradores, bem como a existente na parte externa do prédio que era parte integrante do domicílio. Não foram considerados como cômodo os corredores, alpendres, varandas abertas e também as garagens, depósitos e outros compartimentos utilizados para fins não-residenciais.

Condição de ocupação do domicílio

Quanto à condição de ocupação, classificou-se o domicílio particular permanente, na data de referência, como:

- Próprio, já quitado – quando era de propriedade, total ou parcial, de um ou mais moradores e já estava integralmente pago;

- Próprio, em aquisição – quando era de propriedade, total ou parcial, de um ou mais moradores e ainda não estava integralmente pago;

- Alugado – quando o aluguel era pago por um ou mais moradores. Considerou-se também como alugado o domicílio em que o empregador – de qualquer um dos moradores – pagava, como parte integrante do salário, uma parcela em dinheiro para pagamento do aluguel;

- Cedido por empregador – quando era cedido por empregador – particular ou público – de qualquer um dos moradores, ainda que mediante a uma taxa de ocupação ou conservação – condomínio, gás, luz, etc. – Incluiu-se, neste caso, o domicílio cujo aluguel era pago diretamente pelo empregador de um dos moradores do domicílio;

- Cedido de outra forma – quando era cedido gratuitamente por pessoa que não era moradora ou por instituição que não era empregadora de algum dos moradores, ainda que mediante uma taxa de ocupação – imposto, condomínio, etc. – ou de conservação. Incluiu-se, neste caso, o domicílio cujo aluguel integral era pago, diretamente ou indiretamente, por pessoa não-moradora ou por instituição não-empregadora de algum morador; ou

- Outra condição – quando era ocupado de forma diferente das anteriormente relacionadas. Incluiu-se, neste caso: o domicílio cujo aluguel, pago por morador, referia-se à unidade domiciliar em conjunto com unidade não-residencial – oficina, loja, etc. – quando a família residisse em domicílio localizado em estabelecimento agropecuário arrendado, ou ainda o domicílio ocupado por invasão.

Condição de ocupação do terreno

Quanto à condição de ocupação, o terreno onde se localizava o domicílio particular permanente foi classificado, na data de referência, como:

- Próprio – quando o terreno ou a fração ideal de terreno – para apartamento), onde se localizava o domicílio, fosse de propriedade, total ou parcial, de pelo menos um de seus moradores;

- Cedido – quando o terreno, onde se localizava o domicílio, fosse cedido a, pelo menos, um de seus moradores; e

- Outra condição – quando o terreno onde se localizava o domicílio apresentava situação diferente das anteriormente relacionadas. Incluíram-se nesta situação os casos de invasão de terreno.

Densidade de moradores por cômodo

A densidade de moradores por cômodo resultou da divisão do total de moradores pelo número de cômodos do domicílio particular permanente.

Densidade de moradores por dormitório

A densidade de moradores por dormitório resultou da divisão do total de moradores pelo número de dormitórios do domicílio particular permanente.

Dependência doméstica

Caracterizava dependência doméstica a situação de subordinação dos empregados domésticos e dos agregados em relação à pessoa responsável pelo domicílio ou pela família.

Destino do lixo

O destino do lixo proveniente do domicílio particular permanente foi classificado, na data de referência, como:

Coletado:

- Coletado por serviço de limpeza – quando o lixo do domicílio era coletado diretamente por serviço de empresa pública ou privada; ou

- Colocado em caçamba de serviço de limpeza – quando o lixo do domicílio era depositado em uma caçamba, tanque ou depósito, fora do domicílio, para depois ser coletado por serviço de empresa pública ou privada.

Não coletado:

- Queimado – na propriedade – quando o lixo do domicílio era queimado no terreno ou na propriedade em que se localizava o domicílio;

- Enterrado – na propriedade – quando o lixo do domicílio era enterrado no terreno ou propriedade em que se localizava o domicílio;

- Jogado em terreno baldio ou logradouro – quando o lixo do domicílio era jogado em terreno baldio ou logradouro público;

- Jogado em rio, lago ou mar – quando o lixo do domicílio era jogado nas águas de rio, lago ou mar; ou

- Outro destino – quando o lixo do domicílio tinha destino distinto dos descritos anteriormente.

No caso de existir mais de um destino para o lixo proveniente do domicílio, prevaleceu o que se enquadrava primeiro na ordem relacionada.

Domicílio

Considerou-se domicílio o local estruturalmente separado e independente que, na data de referência, destinava-se a servir de habitação a uma ou mais pessoas, ou que estava sendo utilizado como tal.

Domicílio coletivo

Estabelecimento ou instituição onde a relação entre as pessoas que nele habitavam na data de referência estava restrita a normas de subordinação administrativa. São exemplos de domicílio coletivo: hotéis, pensões, presídios, cadeias, penitenciárias, quartéis, postos militares, asilos, orfanatos, conventos, hospitais e clínicas - com internação - motéis, *campings*, etc.

Domicílio particular

Considerou-se domicílio particular a moradia onde, na data de referência, o relacionamento entre seus ocupantes era ditado por laços de parentesco, por dependência doméstica ou por normas de convivência.

Domicílio particular de uso ocasional

Foi caracterizado como de uso ocasional o domicílio particular permanente que, na data de referência, servia ocasionalmente de moradia, ou seja, era o domicílio usado para descanso de fins de semana, férias ou outro fim, mesmo que seus ocupantes ocasionais estivessem presentes.

Domicílio particular fechado

Domicílio particular permanente que, na data de referência, estava ocupado, porém seus moradores, durante todo o período da coleta, estiveram temporariamente ausentes.

Domicílio particular improvisado

Domicílio localizado em unidade não-residencial – loja, fábrica, etc. – que não tinha dependências destinadas exclusivamente à moradia, mas que, na data de referência, estava ocupado por morador(es).

Também foram considerados como domicílios particulares improvisados os prédios em construção; vagões de trem, carroças, tendas, barracas, *trailers*, grutas, aqueles situados sob pontes, viadutos, etc., que estavam servindo de moradia.

Domicílio particular ocupado

Domicílio particular, permanente ou improvisado, que, na data de referência, estava ocupado por moradores, no qual foi possível realizar a entrevista, durante o período da coleta.

Domicílio particular permanente

Domicílio que foi construído para servir exclusivamente à habitação e, na data de referência, tinha finalidade de servir de moradia a uma ou mais pessoas.

Foram também considerados domicílios particulares permanentes em estabelecimentos institucionais – como hospitais, leprosários, asilos, mosteiros, quartéis, escolas, prisões e assemelhados – aqueles localizados em edificações independentes e que:

- estavam ocupados por famílias cujos membros, um ou mais, eram empregados ou donos do estabelecimento;

- estavam ocupados por famílias cujos membros, um ou mais, faziam parte ou não da instituição, como nos leprosários, colônias correcionais, etc.; e

- estavam ocupados por famílias cujos membros, um ou mais, faziam parte ou não de estabelecimentos ou zonas militares.

Domicílio particular vago

Domicílio particular permanente que não tinha morador na data de referência

Dormitório

Considerou-se como dormitório o quarto ou qualquer outra dependência do domicílio particular permanente, que estivesse servindo de dormitório aos moradores, na data de referência.

Espécie do domicílio

Quanto à espécie, o domicílio foi classificado em: domicílio particular – permanente ou improvisado – ou domicílio coletivo.

Espécie do domicílio particular

Quanto à espécie, o domicílio particular foi classificado em: domicílio particular ocupado, domicílio particular fechado, domicílio particular de uso ocasional, ou domicílio particular vago.

Existência de bens

No domicílio particular permanente pesquisou-se, na data de referência;

- a existência e quantidade de automóveis para uso particular, televisores e aparelhos de ar-condicionado; e

- a existência de qualquer tipo de aparelho de rádio – inclusive a pilha ou acoplado a outros aparelhos – geladeira ou freezer, videocassete, máquina de lavar roupas – exclusive a que não realizava as operações de enxágüe e centrifugação – forno de microondas, linha telefônica instalada - excluindo os aparelhos celulares – e microcomputador.

Foram considerados somente os bens que estivessem em condições de uso, independentemente de serem próprios, cedidos ou alugados.

Forma de abastecimento de água

A forma de abastecimento de água do domicílio particular permanente foi classificada, na data de referência, como:

- Rede geral – quando o domicílio, ou o terreno ou a propriedade em que estava localizado, estava ligado à rede geral de abastecimento de água;

- Poço ou nascente – na propriedade – quando o domicílio era servido por água de poço ou nascente, localizado no terreno ou na propriedade em que estava construído o domicílio; ou

- Outra – quando o domicílio era servido por água de poço ou nascente localizado fora do terreno ou da propriedade em que estava construído ou quando o domicílio era servido de água de reservatório (ou caixa), abastecido com água das chuvas, por carro-pipa ou, ainda, por poço ou nascente localizados fora do terreno ou da propriedade em que estava construído.

Iluminação elétrica

Considerou-se como tendo iluminação elétrica o domicílio particular permanente, que, na data de referência, dispunha desse tipo de iluminação, proveniente ou não de uma rede geral, com ou sem medidor.

Independência

Um local de habitação era dito independente quando possuía acesso direto que permitia aos seus moradores entrar e sair sem necessidade de passar por locais de moradia de outras pessoas.

Normas de convivência

São regras estabelecidas para convivência de pessoas que residiam no mesmo domicílio, que não estavam ligadas por laços de parentesco ou dependência doméstica.

Residência principal - pessoa que ocupava duas ou mais residências

Foi necessário que se estabelecesse, junto à pessoa entrevistada, qual era sua residência principal na data de referência. Foram utilizados os critérios seguintes, na ordem em que estão relacionados, para que a pessoa entrevistada determinasse onde realmente residia:

- a pessoa indicava qual a residência habitual – residência principal ;

- a pessoa era considerada moradora na residência em que passava a maior parte do ano; e

- a pessoa era considerada moradora na residência que tinha há mais tempo, caso ocupasse duas residências em períodos iguais durante o ano.

Sanitário

Para o domicílio particular permanente que, na data de referência, não tinha banheiro, investigou-se a existência de sanitário no domicílio, no terreno ou na propriedade em que se localizava, de uso exclusivo ou não dos moradores.

Considerou-se como sanitário o local, no terreno ou na propriedade onde se localizava o domicílio particular permanente, limitado por paredes de qualquer material, coberto ou não por um teto, que dispunha de aparelho sanitário ou de buraco para dejeções.

Separação

Um local de habitação era dito separado quando fosse limitado por paredes, muros ou cercas, coberto por um teto e permitisse que uma ou mais pessoas, que nele habitavam, isolassem-se das demais, com a finalidade de dormir, preparar e/ou consumir seus alimentos e proteger-se do meio ambiente, arcando, total ou parcialmente, com suas despesas de alimentação ou moradia.

Situação do domicílio

Segundo a localização do domicílio, a situação era urbana ou rural, conforme definida por lei municipal em vigor em 1º de agosto de 2000. Na situação urbana foram consideradas as áreas urbanizadas ou não, correspondentes às cidades (sedes municipais), às vilas (sedes distritais) ou às áreas urbanas isoladas. A situação rural abrangeu toda a área situada fora desses limites, inclusive os aglomerados rurais de extensão urbana, os povoados e os núcleos.

Tipo de esgotamento sanitário

O tipo de esgotamento sanitário do banheiro ou sanitário do domicílio particular permanente foi classificado, na data de referência, como:

- Rede geral de esgoto ou pluvial – quando a canalização das águas servidas e dos dejetos provenientes do banheiro ou sanitário estava ligada a um sistema de coleta que os conduzia a um escoadouro geral da área, região ou município, mesmo que o sistema não dispusesse de estação de tratamento da matéria esgotada;

- Fossa séptica – quando a canalização do banheiro ou sanitário estava ligada a uma fossa séptica, ou seja, a matéria era esgotada para uma fossa próxima, onde passava por um processo de tratamento ou decantação, sendo, ou não, a parte líquida conduzida em seguida para um desaguadouro geral da área, região ou município;

- Fossa rudimentar – quando o banheiro ou sanitário estava ligado a uma fossa rústica – fossa negra, poço, buraco, etc.;

- Vala - quando o banheiro ou sanitário estava ligado diretamente a uma vala a céu aberto;

- Rio, lago ou mar – quando o banheiro ou sanitário estava ligado diretamente a um rio, lago ou mar; ou

- Outro escoadouro – quando o escoadouro dos dejetos provenientes do banheiro ou sanitário fosse diferente do descrito anteriormente.

No caso de existir mais de um tipo de escoadouro, prevaleceu o que se enquadrava primeiro na ordem relacionada.

Tipo do domicílio

Quanto ao tipo, classificou-se, na data de referência, o domicílio particular permanente como:

- Casa – quando localizado em uma edificação de um ou mais pavimentos, desde que ocupado integralmente por um único domicílio, com acesso direto a um logradouro – arruamento, vila, avenida, caminho, etc. – legalizado ou não, independentemente do material utilizado em sua construção;

- Apartamento – quando localizado em edifício de um ou mais andares, com mais de um domicílio, servido por espaços comuns – hall de entrada, escadas, corredores, portaria ou outras dependências. Inclui-se, neste caso, o domicílio localizado em um prédio de dois ou mais andares em que as demais unidades não eram residenciais e, ainda, aquele localizado em edifício de dois ou mais pavimentos com entradas independentes para os andares; ou

- Cômodo – quando composto por um ou mais aposentos localizados em uma casa de cômodos, cortiço, cabeça-de-porco, etc.

Unidade não-residencial

Unidade destinada a atividades comerciais, industriais ou de prestação de serviços, onde não existiam moradores na data de referência.

As embaixadas, legações – representações diplomáticas de categoria inferior a embaixadas – ou consulados, por representarem território estrangeiro, foram considerados unidades não-residenciais, ainda que estivessem servindo de moradia.

6.3.5 Família e pessoa

Os conceitos e formas de investigação, referentes à investigação das características da pessoa e da família, são apresentados a seguir, em ordem alfabética.

Abono de permanência em serviço

Rendimento recebido no mês de referência, a quem permanecia trabalhando, embora tivesse tempo de serviço suficiente para se aposentar.

Alfabetização

Considerou-se como alfabetizada a pessoa capaz de ler e escrever um bilhete simples no idioma que conhecia. Aquela que aprendeu a ler e escrever, mas esqueceu, e a que apenas assinava o próprio nome foram consideradas analfabetas.

Aluguel

Rendimento recebido no mês de referência, a título de locação, sublocação ou arrendamento de móveis, imóveis, máquinas, equipamentos, animais, etc.

Anos de estudo

Essa classificação foi obtida, em função da última série concluída com aprovação, no nível ou grau mais elevado que a pessoa de 5 anos ou mais de idade, na data de referência, estava freqüentando ou havia freqüentado, ou seja:

- Sem instrução e menos de um ano de estudo – para a pessoa que nunca freqüentou escola ou, embora tenha freqüentado, não concluiu pelo menos a 1ª série do ensino fundamental, 1º grau ou elementar;

- um ano de estudo – para a pessoa que concluiu curso de alfabetização de adultos; ou a 1ª série do ensino fundamental, 1º grau ou elementar;

- dois anos de estudo – para a pessoa que concluiu a 2ª série do ensino fundamental, 1º grau ou elementar;

- três anos de estudo – para a pessoa que concluiu a 3ª série do ensino fundamental, 1º grau ou elementar;

- quatro anos de estudo – para a pessoa que concluiu a 4ª série do ensino fundamental ou 1º grau; ou, no mínimo, a 4ª série e, no máximo, a 6ª série do elementar;

- cinco anos de estudo – para a pessoa que concluiu a 5ª série do ensino fundamental ou 1º grau; ou a 1ª série do médio 1º ciclo;

- seis anos de estudo – para a pessoa que concluiu a 6ª série do ensino fundamental ou 1º grau; ou a 2ª série do médio 1º ciclo;

- sete anos de estudo – para a pessoa que concluiu a 7ª série do ensino fundamental ou 1º grau; ou a 3ª série do médio 1º ciclo;

- oito anos de estudo – para a pessoa que concluiu a 8ª série do ensino fundamental ou 1º grau; ou, no mínimo, a 4ª série e, no máximo, a 5ª série do médio 1º ciclo;

- nove anos de estudo – para a pessoa que concluiu a 1ª série do ensino médio, 2º grau ou médio 2º ciclo;

- dez anos de estudo – para a pessoa que concluiu a 2ª série do ensino médio, 2º grau ou médio 2º ciclo;

- onze anos de estudo – para a pessoa que concluiu, no mínimo, a 3ª série e, no máximo, a 4ª série do ensino médio, 2º grau ou médio 2ª ciclo;

- doze anos de estudo – para a pessoa que concluiu a 1ª série do superior;

- treze anos de estudo – para a pessoa que concluiu a 2ª série do superior;

- quatorze anos de estudo – para a pessoa que concluiu a 3ª série do superior;

- quinze anos de estudo – para a pessoa que concluiu a 4ª série do superior;

- dezesseis anos de estudo – para a pessoa que concluiu a 5ª série do superior;

- dezessete anos de estudo ou mais – para a pessoa que concluiu a 6ª série do superior ou mestrado ou doutorado; ou

- Não determinados – para a pessoa que freqüentava ensino fundamental ou 1º grau não-seriado.

Aposentadoria

Rendimento recebido de instituto de previdência oficial no mês de referência, a título de aposentadoria, jubilação ou reforma, incluindo, se for o caso, a parcela paga por entidade de previdência privada fechada como complementação de aposentadoria.

Atividade

Investigou-se, na semana de referência, a atividade correspondente ao trabalho principal da pessoa. Para isso, foi considerada a principal finalidade ou o ramo do negócio, da firma, da instituição, da empresa ou da entidade a que a pessoa prestava serviços ou, para a que trabalhava por conta própria, a natureza da atividade exercida.

Para a pessoa que trabalhava:

- ligada às atividades de agricultura, silvicultura, pecuária, extração vegetal ou mineral, pesca e piscicultura, considerou-se a atividade principal do estabelecimento;

- em empreendimento agroindustrial, foi considerada a atividade principal, quando se tratava de pessoa que explorava o empreendimento, ou a atividade na qual exercia sua ocupação, quando se tratava de qualquer outro trabalhador;

- em empreendimento que alocava seu pessoal para prestar serviços em outros empreendimentos, foi considerada a atividade principal daquele com o qual possuía vínculo de trabalho; e

- na fabricação de determinados itens – rótulos, embalagens, vidros, etc. – unicamente para adicioná-los ao seu produto final, foi considerada a atividade referente à fabricação deste produto final.

Bolsa-escola

Rendimento recebido de órgão governamental no mês de referência, com vistas a garantir a permanência na escola, das crianças de uma família que estivessem em idade de freqüência obrigatória à escola.

Brasileiro nato

Pessoa que nasceu no Brasil ou nasceu em país estrangeiro e foi registrada como brasileira, segundo as leis do Brasil.

Categoria do emprego

Os empregados foram classificados, quanto à categoria do emprego no trabalho principal, em: com carteira de trabalho assinada; militares e funcionários públicos estatutários; e outros sem carteira de trabalho assinada.

A categoria dos militares e funcionários públicos estatutários foi constituída pelos militares das Forças Armadas – Exército, Marinha e Aeronáutica – ou das Forças Auxiliares – Polícia Militar, Corpo de Bombeiros Militar, etc. 0 e pelos empregados de instituição, fundação, autarquia, etc., do Poder Público – Executivo, Legislativo ou Judiciário – regidos pelo Regime Jurídico dos Funcionários Públicos de qualquer instância – federal, estadual e municipal).

Condição de atividade

Quanto à condição de atividade, a pessoa foi classificada na semana de referência, como economicamente ativa ou não-economicamente ativa.

Condição na família e Condição no domicílio

Cada morador foi classificado, na data de referência, segundo os seguintes critérios:

- Pessoa responsável – para o homem ou a mulher responsável pelo domicílio ou pela família ou que assim era considerado(a) pelos demais moradores.

- Cônjuge, companheiro(a) – para o homem ou a mulher que vivia conjugalmente com a pessoa responsável pelo domicílio ou pela família, existindo ou não vínculo matrimonial.

- Filho(a), enteado(a) – para o(a) filho(a) ou enteado(a), inclusive o(a) filho(a) adotivo(a) ou de criação, da pessoa responsável pelo domicílio ou pela família e, também, para o(a) filho(a) somente do(a) cônjuge, mesmo sendo o(a) cônjuge já falecido(a) ou não morando mais no domicílio.

- Pai, mãe, sogro(a) – para o pai, a mãe ou o(a) sogro(a), inclusive o padrasto ou a madrasta, da pessoa responsável pelo domicílio ou pela família.

- Neto(a), bisneto(a) – para o(a) neto(a) ou bisneto(a) da pessoa responsável pelo domicílio ou pela família, inclusive para o(a) neto(a) ou bisneto(a) somente do cônjuge.

- Irmão ou irmã – para o irmão ou a irmã da pessoa responsável pelo domicílio ou pela família, inclusive os que não tinham laços consangüíneos (adotivos ou de criação).

- Outro parente – para o(a) avô(ó), o(a) bisavô(ó), o genro, a nora, o(a) cunhado(a), o(a) tio(a), o(a) sobrinho(a) ou o(a) primo(a) da pessoa responsável pelo domicílio ou pela família, inclusive para a pessoa que tinha um destes laços de parentesco somente com o cônjuge.

- Agregado – para a pessoa que, sem ser parente, pensionista, empregado doméstico ou parente do empregado doméstico, não pagava hospedagem nem contribuía para as despesas de alimentação e moradia no domicílio.

- Pensionista – para a pessoa que, sem ser parente, pagava hospedagem ou contribuía para as despesas de alimentação e moradia no domicílio.

- Empregado(a) doméstico(a) – para a pessoa que prestava serviços domésticos remunerados a um ou mais membros do domicílio.

- Parente do(a) empregado(a) doméstico(a) – para a pessoa que era parente do(a) empregado(a) doméstico(a) e que não prestava serviços domésticos remunerados a membro do domicílio.

- Individual em domicílio coletivo – para a pessoa só que residia em domicílio coletivo, ainda que compartilhando a unidade com outra(s) pessoa(s) com a(s) qual(is) não tinha laços consangüíneos.

Contribuição para instituto de previdência oficial

Foi pesquisado se a pessoa contribuía para um dos seguintes institutos de previdência oficial no trabalho principal:

- Instituto Nacional de Seguridade Social – INSS;

- Plano de Seguro Social da União; ou

- Instituto de Previdência Social Estadual ou Municipal.

Coorte

Conjunto de indivíduos que estão experimentando um acontecimento similar no transcurso de um mesmo período de tempo.

Coorte hipotética de mulheres

Num censo demográfico, a classificação das mulheres por grupos qüinqüenais de idade, dentro do período fértil, está associada a uma análise de período. Uma análise de coorte considera, por exemplo, um grupo de mulheres que ingressa no período fértil e, ao longo do tempo, observa-se o comportamento do mesmo frente aos riscos de procriação. Entretanto, em um único censo demográfico, mesclam-se distintas gerações de mulheres e, de acordo com o conceito da Taxa de Fecundidade Total, supõe-se o acompanhamento de como essas mulheres vão tendo seus filhos ao longo do tempo. Por esse motivo, na definição conceitual da Taxa de Fecundidade Total, é necessário enfatizar que o grupo de mulheres em questão trata-se de uma coorte hipotética.

Cor ou raça

A investigação da cor ou raça ocorreu através da classificação, pela própria pessoa, em uma das seguintes opções:

- Branca – para a pessoa que se declarou branca;

- Preta – para a pessoa que se declarou preta;

- Amarela – para a pessoa que se declarou de raça amarela de origem japonesa, chinesa, coreana, etc.;

- Parda – para a pessoa que se declarou parda, mulata, cabocla, cafuza, mameluca ou mestiça; ou

- Indígena – para a pessoa que se declarou indígena ou índia.

Data de nascimento ou idade presumida do último filho tido nascido vivo até 31 de julho de 2000

Registrou-se o mês e o ano de nascimento do último filho tido nascido vivo até 31 de julho de 2000. Se, esgotados todos os esforços, não foi possível a obtenção do mês e ano de nascimento do último filho tido nascido vivo, registrou-se sua idade presumida, fornecida pela pessoa entrevistada.

Deficiência

Foi pesquisado se a pessoa era portadora de deficiência mental permanente e, ainda, de um dos seguintes tipos de deficiência física permanente: tetraplegia, paraplegia, hemiplegia, falta de membro ou de parte dele.

Foi investigado, também, se a pessoa era portadora de deficiência auditiva, visual ou motora, por meio da avaliação, feita pela própria pessoa, do seu grau de incapacidade – incapaz, com grande dificuldade ou com alguma dificuldade permanente, sem nenhuma dificuldade – de enxergar, ouvir e caminhar ou subir escada.

Deficiência auditiva

De acordo com a avaliação do grau de incapacidade auditiva, – feita pela pessoa, levando em conta o uso de aparelho auditivo, no caso de utilizá-lo) a classificação foi:

- Incapaz – quando a pessoa se declarou totalmente surda;

- Grande dificuldade permanente – quando a pessoa declarou ter grande dificuldade permanente de ouvir, ainda que usando aparelho auditivo;

- Alguma dificuldade permanente – quando a pessoa declarou ter alguma dificuldade permanente de ouvir, ainda que usando aparelho auditivo; ou

- Nenhuma dificuldade – quando a pessoa declarou sem dificuldade para ouvir, ainda que isto exigisse o uso de aparelho auditivo.

Deficiência física

Foram investigadas as seguintes deficiências:

- Tetraplegia – a paralisia permanente total de ambos os braços e pernas - quadriplegia;

- Paraplegia – a paralisia permanente das pernas;

- Hemiplegia – a paralisia permanente de um dos lados do corpo; ou

- Falta de membro ou de parte dele – a falta de perna, braço, mão, pé ou do dedo polegar ou a falta de parte da perna ou braço.

Deficiência mental permanente

Retardamento mental resultante de lesão ou síndrome irreversível, que se caracterizava por dificuldades ou limitações intelectuais associadas a duas ou mais áreas de habilidades adaptativas, tais como: comunicação, cuidado pessoal, autodeterminação, cuidados com saúde e segurança, aprendizagem, lazer, trabalho, etc.

Não se consideraram como deficiência mental as perturbações ou doenças mentais como autismo, neurose, esquizofrenia e psicose.

Deficiência motora

De acordo com a avaliação do grau de incapacidade de caminhar/subir escadas – feita pela pessoa, levando em conta o uso de prótese, bengala, ou aparelho auxiliar, no caso de utilizá-los – a classificação foi:

- Incapaz – quando a pessoa se declarou incapaz de caminhar ou subir escadas, sem ajuda de outra pessoa;

- Grande dificuldade permanente – quando a pessoa declarou ter grande dificuldade permanente de caminhar ou subir escadas, sem ajuda de outra pessoa, ainda que usando prótese, bengala ou aparelho auxiliar;

- Alguma dificuldade permanente – quando a pessoa declarou ter alguma dificuldade permanente de caminhar ou subir escadas, sem ajuda de outra pessoa, ainda que usando prótese, bengala ou aparelho auxiliar; ou

- Nenhuma dificuldade – quando a pessoa declarou não ter qualquer dificuldade permanente de caminhar ou subir escadas, sem ajuda de outra pessoa, ainda que usando prótese, bengala ou aparelho auxiliar; foram também assim classificadas as crianças que, por pouca idade, apresentavam pouca ou nenhuma capacidade de caminhar.

Deficiência visual

De acordo com a avaliação do grau de incapacidade visual – feita pela pessoa, levando em conta o uso de óculos ou lentes de contato, no caso de utilizá-los – a classificação foi:

- Incapaz de enxergar – quando a pessoa se declarou totalmente cega;

- Grande dificuldade permanente de enxergar – quando a pessoa declarou ter grande dificuldade permanente de enxergar, ainda que usando óculos ou lentes de contato;

- Alguma dificuldade permanente de enxergar – quando a pessoa declarou ter alguma dificuldade de enxergar, ainda que usando óculos ou lentes de contato; ou

- Nenhuma dificuldade – quando a pessoa declarou sem dificuldade para enxergar, ainda que isto exigisse o uso de óculos ou lentes de contato.

Deslocamento para trabalho ou estudo

Foi investigado se a pessoa recenseada trabalhava ou estudava no município de residência, na data de referência; se não trabalhava nem estudava ou se estudava ou trabalhava em outro município ou país estrangeiro. A investigação considerou os conceitos de trabalho e freqüência à escola descritos neste texto.

Educação

Foram investigados para a pessoa que freqüentava escola, na data de referência, a alfabetização, a rede de ensino, o tipo de curso e a série freqüentados. Para a que não estava freqüentando, mas anteriormente havia freqüentado escola, pesquisou-se o curso mais elevado, a última série concluída com aprovação, a conclusão ou não do curso e, para quem tinha curso superior, a espécie do curso mais elevado concluído.

Empreendimento

Definiu-se como empreendimento a empresa, a instituição, a entidade, a firma, o negócio, etc., ou, ainda, o trabalho sem estabelecimento, desenvolvido individualmente ou com ajuda de outras pessoas – empregados, sócios ou trabalhadores não-remunerados.

Estado Civil

O estado civil da pessoa foi investigado, na semana de referência, da seguinte forma:

- Casado(a) – quando a pessoa tinha o estado civil de casada;

- Desquitado (a) ou separado (a) judicialmente – quando a pessoa tinha o estado civil de desquitada ou separada homologado por decisão judicial;

- Divorciado(a) – quando a pessoa tinha o estado civil de divorciada homologado por decisão judicial;

- Viúvo (a) – quando a pessoa tinha o estado civil de viúva; ou

- Solteiro(a) – quando a pessoa tinha o estado civil de solteira.

Estrangeiro

Pessoa que nasceu fora do Brasil ou aquela que nasceu no Brasil e se registrou em representação estrangeira, e não se naturalizou brasileira.

Família

Nos domicílios particulares, considerou-se como família:

- a pessoa que morava sozinha;

- o conjunto de pessoas ligadas por laços de parentesco ou por dependência doméstica; e

- as pessoas ligadas por normas de convivência.

Nos domicílios coletivos, considerou-se como família apenas as pessoas que tinham laços de parentesco ou dependência doméstica.

Famílias conviventes

Foram definidas como conviventes as famílias com, no mínimo, duas pessoas cada uma, que conviviam no mesmo domicílio particular, na data de referência.

Fecundidade

A investigação de fecundidade abrangeu as mulheres de 10 anos ou mais de idade, na data de referência, sendo investigados os seguintes aspectos:

- o número de filhos tidos até 31 de julho de 2000, segundo o sexo, identificando os nascidos vivos, os nascidos mortos e os que permaneciam vivos;

- o sexo do último filho nascido vivo, até 31 de julho de 2000; e

- o mês e o ano de nascimento ou, na falta de um destes, a idade presumida e o sexo do (a) último (a) filho (a) nascido (a) vivo (a), bem como se permanecia vivo até 31 de julho de 2000.

Filho tido nascido vivo

Considerou-se como filho tido nascido vivo aquele que, após a expulsão ou extração completa do corpo da mãe, independentemente do tempo de duração da gravidez, manifestou qualquer sinal de vida – espiração, choro, movimento de músculos de contração voluntária, batimento cardíaco, etc. ainda que tenha falecido em seguida.

Filho nascido morto

Considerou-se como filho nascido morto o resultado de gestação igual ou superior a 7 (sete) meses e que, após a separação do corpo da mãe, não tinha mostrado qualquer evidência de vida, tais como: respiração, batimento cardíaco, pulsação do cordão umbilical, etc.

Freqüência à creche

Considerou-se como freqüentando creche a criança que, na data de referência, estava em estabelecimento destinado a dar assistência diurna às crianças nas primeiras idades.

Freqüência à escola

Considerou-se como freqüentando escola a pessoa matriculada, na data de referência, em:

- curso regular – pré-escolar, ensino fundamental, 1º grau, ensino médio, 2º grau, superior (3º grau), mestrado ou doutorado;

- curso supletivo – ensino fundamental (1º grau), ensino médio (ou 2º grau), ministrado em escola;

- curso de alfabetização de adultos; e

- curso pré-vestibular.

Horas trabalhadas por semana

Pesquisou-se o número de horas inteiras – desprezando-se ou considerando-se como uma hora, respectivamente, os períodos inferiores ou superiores a 30 minutos – habitualmente trabalhadas por semana, separadamente, no trabalho principal e nos demais trabalhos, que a pessoa tinha na semana de referência.

Levou-se em conta, também, as horas que a pessoa ocupava habitualmente, fora do local de trabalho, com tarefas relacionadas com sua ocupação, sendo, por outro lado, excluídos os períodos destinados normalmente às refeições.

Idade

A investigação da idade da pessoa foi feita da seguinte forma:

- pela pesquisa: do mês e ano do nascimento;

- pela declaração da idade que tinha em 31 de julho de 2000, somente para quem respondeu o mês e ano de nascimento; nesse caso, foi pesquisado o número de anos ou meses completos, respectivamente, para as pessoas com idades superiores ou inferiores a um ano; para o recém-nascido, com menos de um mês de vida, considerou-se a idade zero; e

- pela idade presumida, para a pessoa que não sabia o mês ou o ano do nascimento.

Assim, para efeito da divulgação, a idade do entrevistado será a calculada através do mês e ano de nascimento, ou a idade presumida. A idade em anos completos foi pesquisada tendo em vista a realização de estudos específicos e também para apoio às atividades de crítica dos dados.

Índice de Gini

O Índice de Gini é uma medida do grau de concentração de uma distribuição, cujo valor varia de zero (a perfeita igualdade) até um (a desigualdade máxima).

O índice de Gini da distribuição do rendimento é calculado através da seguinte expressão:

$$G = \frac{\sum_{i=1}^{N}(2i-1)y_{(i)}}{N\sum_{i=1}^{N} y_i} - 1$$

onde: y_i é o valor da variável y para a i-ésima unidade da população, e $y_{(i)}$ é o valor da i-ésima estatística de ordem da variável y na população.

Um estimador consistente para G é dado por Sandström, Wretman e Waldén (1988):

$$\hat{G}_w = \frac{\sum_{i=1}^{n}\left[2\hat{F}(y_{(i)}) + w_{(i)}\right]w_{(i)}y_{(i)}}{\hat{N}\sum_{i=1}^{n} w_i y_i} - 1$$

onde:

w_i é o peso da i-ésima unidade amostral, satisfazendo a seguinte propriedade

$E\left[\hat{N}\right] = E\left[\sum_{i=1}^{n} w_i\right] \cong N$ (pesos devem permitir obter estimadores aproximadamente não viciados de total);

$\hat{F}(t) = \sum_{j=1}^{n} w_j I(y_j < t)$ para qualquer $t \in \Re$;

$I(y_j < t)$ é a função indicadora do evento $y_j < t$, isto é, a função que toma o valor 1 quando o evento ocorre e valor zero quando não ocorre;

$y_{(i)}$ é o valor da i-ésima estatística de ordem da variável y na amostra;

$w_{(i)}$ é o peso correspondente à i-ésima estatística de ordem $y_{(i)}$.

Lugar de nascimento

Para os brasileiros natos, foi pesquisada a Unidade da Federação e, para os naturalizados brasileiros e estrangeiros, o país de nascimento.

Lugar de residência anterior

Para as pessoas que, na data de referência, residiam há menos de 10 anos na Unidade da Federação, investigou-se a unidade da federação ou país estrangeiro de residência anterior.

Lugar de residência em 31 de julho de 1995

Pesquisou-se o lugar – município e Unidade da Federação ou país estrangeiro – em que a pessoa residia em 31 de julho de 1995.

Mesada

Rendimento recebido no mês de referência, em caráter regular, sem contrapartida de serviços prestados, proveniente de pessoa não-moradora do domicílio.

Migração

Foram investigados o lugar de nascimento, o tempo de residência, o lugar de residência anterior, o lugar de residência em 31 de julho de 1995 e o deslocamento para trabalho ou estudo.

Morador

Considerou-se como moradora a pessoa que:

- tinha o domicílio como local de residência habitual e nele se encontrava, na data de referência; ou

- embora ausente na data de referência, tinha o domicílio como residência habitual, desde que esta ausência não fosse superior a 12 meses, em relação àquela data, por um dos seguintes motivos:

- viagens: a passeio, a serviço, a negócio, de estudos, etc.;

- internação em estabelecimento de ensino ou hospedagem em outro

domicílio, visando a facilitar a freqüência à escola durante o ano letivo;

- detenção sem sentença definitiva declarada;

-internação temporária em hospital ou estabelecimento similar; e

- embarque a serviço (marítimos).

Foram consideradas moradoras nos locais em que se encontravam na data de referência, independentemente do período de afastamento do domicílio, as pessoas nas seguintes condições:

- internadas permanentemente em sanatórios, asilos, conventos ou estabelecimentos similares;

- moradoras em pensionatos e que não tinham outro local habitual de residência;

- condenadas com sentença definitiva declarada; e

- que migraram para outras regiões, em busca de trabalho, e ali fixaram residência.

Nacionalidade

A nacionalidade das pessoas foi investigada como: brasileiro nato, naturalizado brasileiro e estrangeiro.

Naturalizado brasileiro

Pessoa que nasceu em país estrangeiro e obteve a nacionalidade brasileira por meio de título de naturalização ou valendo-se de disposição da legislação brasileira.

Natureza da união

A natureza da última união conjugal da pessoa foi investigada como:

- Casamento civil e religioso – quando a pessoa vivia ou já tinha vivido em companhia de cônjuge, com quem era casada no civil e no religioso, inclusive a pessoa que, embora somente tenha comparecido à cerimônia religiosa, regularizou o ato civil de acordo com a legislação vigente;

- Só casamento civil – quando a pessoa vivia ou já tinha vivido em companhia de cônjuge com quem era casada somente no civil;

- Só no religioso – quando a pessoa vivia ou já tinha vivido em companhia de cônjuge com quem era casada somente no religioso, em qualquer religião ou culto;

- União consensual – quando a pessoa vivia ou já tinha vivido em companhia de cônjuge com quem não contraiu casamento civil ou religioso; ou

- Nunca viveu – quando a pessoa nunca tinha vivido em companhia de cônjuge ou companheiro (a).

Nível de ensino

Para a pessoa que freqüentava escola ou creche, o seu nível de ensino foi classificado através do tipo de curso que freqüentava, da seguinte maneira: creche; pré-escolar ou classe de alfabetização; alfabetização de adultos; fundamental, médio, pré-vestibular; superior de graduação; e mestrado ou doutorado.

Número de empregados

Para as pessoas classificadas, na semana de referência, como empregadores, foi investigado o número de pessoas empregadas no empreendimento em que trabalhavam, no mês de referência, exclusive os sócios e trabalhadores

não-remunerados, obedecendo ao seguinte: uma pessoa empregada; duas pessoas empregadas; de três a cinco pessoas empregadas; de seis a dez pessoas empregadas; e onze ou mais pessoas empregadas.

Nos empreendimentos que possuíam mais de um estabelecimento, foi investigado o total de pessoas empregadas em todos os estabelecimentos.

Número de trabalhos

Pesquisou-se o número de trabalhos, ou seja, em quantos empreendimentos a pessoa teve trabalho na semana de referência.

O trabalho na produção para o próprio consumo somente seria contado se a pessoa não tivesse qualquer outro trabalho remunerado ou sem remuneração na semana de referência.

Por convenção, o exercício do serviço doméstico remunerado, independentemente do número de unidades domiciliares em que este serviço era prestado, foi contado como se fosse um único trabalho.

Nupcialidade

Para a pessoa de 10 anos ou mais de idade na data de referência, foi investigada a união conjugal, o estado civil e, também, a natureza da última união.

Ocupação

Investigou-se, na semana de referência, a ocupação da pessoa no trabalho principal. Foi considerada como ocupação a função, cargo, profissão ou ofício desempenhado numa atividade econômica.

Pensão

Rendimento recebido de instituto de previdência oficial no mês de referência, deixado por alguém da qual a pessoa era beneficiária.

Pensão alimentícia

Rendimento recebido no mês de referência, a título de pensão alimentícia, paga espontaneamente ou judicialmente.

Pessoa desocupada

Foi classificada como desocupada na semana de referência, a pessoa sem trabalho nessa semana que tomou alguma providência efetiva na procura de trabalho, no período de referência de 30 dias.

Pessoa economicamente ativa

Foi classificada como economicamente ativa na semana de referência, a pessoa ocupada ou desocupada nessa semana.

Pessoa não-economicamente ativa

Foi classificada como não-economicamente ativa na semana de referência, a pessoa que não era ocupada nem desocupada nessa semana.

Pessoa ocupada

Foi classificada como ocupada na semana de referência, a pessoa que tinha trabalho durante todo ou parte desse período. Considerou-se ainda como ocupada a pessoa que não exerceu o trabalho remunerado que tinha na semana de referência por motivo de férias, licença, greve, etc.

População residente

A população residente é formada pelos moradores em domicílios, na data de referência.

Posição na ocupação

Investigou-se a posição na ocupação correspondente ao trabalho principal. Para isso, considerou-se a relação de trabalho existente entre a pessoa e o empreendimento em que trabalhava, podendo ser:

- Trabalhador doméstico – pessoa que prestava serviços domésticos remunerados, em um ou mais domicílios;

- Empregado – pessoa que trabalhava para um empregador – essoa física ou jurídica – geralmente obrigando-se ao cumprimento de uma jornada de trabalho e recebendo, em contrapartida, uma remuneração em dinheiro, mercadorias, produtos ou benefícios, como: moradia, alimentação, roupas, etc. Incluiu-se também nessa situação:

  • a pessoa que prestava o serviço militar obrigatório;

  • sacerdote, ministro de igreja, pastor, rabino, frade, freira e outros religiosos; e

  • aprendiz ou estagiário remunerado;

- Empregador – pessoa que, na semana de referência, trabalhava explorando o seu próprio empreendimento com pelo menos um empregado;

- Conta própria – pessoa que, na semana de referência, trabalhava explorando o seu próprio empreendimento, sozinha ou com sócio, sem ter empregado, ainda que contando com ajuda de trabalhador não-remunerado;

- Não-remunerado em ajuda a membro do domicílio – Pessoa que trabalhou sem remuneração, durante pelo menos uma hora na semana de referência, em ajuda a membro do domicílio que era empregado na produção de bens primários (que compreende as atividades de agricultura, pecuária, silvicultura, extração vegetal ou mineral, caça, pesca ou piscicultura); ou conta própria ou empregador em qualquer atividade; e

- Trabalhador na produção para o próprio consumo – para pessoa que trabalhou, durante pelo menos uma hora na semana de referência, na produção de bens do ramo que compreende as atividades de agricultura, silvicultura, pecuária, extração vegetal, pesca ou piscicultura, destinados exclusivamente à alimentação de pelo menos um morador do domicílio.

Ainda na investigação da posição na ocupação, foi pesquisado, para os trabalhadores domésticos e empregados, se possuíam carteira de trabalho assinada.

Aprendiz ou estagiário sem remuneração

Pessoa que exerceu uma ocupação sem remuneração, durante pelo menos uma hora na semana de referência, na qualidade de aprendiz de uma profissão ou ofício ou em cumprimento de estágio obrigatório ou voluntário ligado a uma profissão.

Procura de trabalho

Definiu-se como procura de trabalho a tomada de alguma providência para conseguir trabalho, no período de referência de 30 dias, tal como: consultar empregador; participar ou inscrever-se em concurso; consultar agência de empregos ou sindicato; colocar ou responder anúncio; consultar parente, amigo ou colega; tomar providência para iniciar negócio próprio; ou outra providência qualquer que efetivamente tivesse como objetivo conseguir trabalho.

Rede de ensino

A creche ou escola que a pessoa freqüentava foi classificada, de acordo com a rede de ensino a que pertencia, como de:

- Rede particular – quando a creche ou escola era da rede particular, inclusive aquela mantida por associação de moradores, empresa, fundação ou doações; ou

- Rede pública – quando a creche ou escola era da rede pública (federal, estadual ou municipal).

Religião ou culto

Pesquisou-se a seita, culto ou o ramo da religião professada pela pessoa; para a que não professava qualquer religião, a classificação foi "sem religião"; e para a criança que não tinha condição de prestar a informação, foi considerada a religião da mãe.

Remuneração bruta

Entendeu-se por remuneração bruta o pagamento do empregado ou do trabalhador doméstico, sem excluir o salário-família e os descontos correspondentes ao instituto de previdência oficial, imposto de renda, faltas, etc., e sem incluir o 13º salário, a gratificação de férias e a participação de lucros pagos por empregador.

Renda mínima

Rendimento recebido de órgão governamental no mês de referência, com vistas a garantir um rendimento mínimo necessário à sobrevivência de uma família.

Rendimento Mediano

Valor do rendimento que ocupa o ponto central de uma distribuição ordenada dos rendimentos.

Rendimento Médio

Corresponde ao valor resultante da divisão entre a soma dos rendimentos de um conjunto de pessoas, e o número de pessoas desse conjunto.

Rendimento nominal mensal

Na investigação através do Questionário Básico, o rendimento nominal mensal somente foi pesquisado para a pessoa responsável pelo domicílio, de forma direta, ou seja, através do valor global. No caso da investigação pelo Questionário da Amostra, considerou-se a soma do rendimento nominal mensal de trabalho com os provenientes de outras fontes, para a pessoa de 10 anos ou mais de idade.

Rendimento nominal mensal de outras fontes

Pesquisou-se o rendimento nominal mensal de aposentadoria, pensão, aluguel, pensão alimentícia, mesada ou doação recebida de não-morador, renda mínima, bolsa-escola ou seguro-desemprego, relativo ao mês de referência e, também, investigou-se o rendimento habitualmente recebido a título de abono de permanência em serviço, pensão paga integralmente por seguradora ou entidade de previdência privada aberta, juros de aplicações financeiras, dividendos, etc.

Rendimento nominal mensal de trabalho

Pesquisou-se, separadamente, o rendimento nominal mensal do trabalho principal e dos demais trabalhos que a pessoa tinha na semana de referência.

Para a pessoa que recebia rendimento fixo, considerou-se a remuneração bruta (do empregado ou trabalhador doméstico) ou a retirada (do trabalhador por conta própria ou empregador) do mês de referência, ou a que receberia se houvesse trabalhado o mês completo.

Para a pessoa que recebia rendimento variável, considerou-se o valor da remuneração bruta ou da retirada do mês de referência.

Para a pessoa licenciada por instituto de previdência oficial, considerou-se o rendimento bruto do mês de referência, recebido como benefício (auxílio-doença, auxílio por acidente de trabalho, etc.).

Para a pessoa que recebia em produtos ou mercadorias de atividade do ramo que compreende a agricultura, silvicultura, pecuária, extração vegetal ou mineral, pesca e piscicultura, considerou-se o valor real ou estimado, recebido no mês de referência, ou em média mensal, no caso de produção sazonal.

Não foi computado como rendimento do trabalho o valor da produção para o próprio consumo, nem a parcela do pagamento efetuada em benefícios, tais como: moradia, alimentação (refeições, cesta de alimentos, vale ou tíquete-alimentação), vale ou tíquete-transporte, roupas, etc.

Rendimento nominal mensal domiciliar

Considerou-se como rendimento nominal mensal domiciliar a soma dos rendimentos nominais mensais dos moradores do domicílio, exclusive os das pessoas cuja condição no domicílio fosse pensionista, empregado(a) doméstico(a) ou parente do empregado(a) doméstico(a).

Retirada

Entendeu-se por retirada o ganho do trabalhador por conta própria ou do empregador com o seu empreendimento, ou seja, o rendimento bruto menos os gastos (pagamento de empregados, compra de equipamentos, matéria-prima, energia elétrica, telefone, etc.) do empreendimento.

Salário mínimo

Para a apuração dos rendimentos, segundo as classes de salário mínimo, considerou-se o valor que vigorava no mês de referência, que era de R$151,00.

Seguro-desemprego

Rendimento recebido no mês de referência, por dispensa de emprego com carteira de trabalho assinada desde que fossem atendidos os requisitos necessários para o recebimento deste benefício.

Taxa de fecundidade total

Expressa o número de filhos que, em média, teria uma mulher, pertencente a uma coorte hipotética de mulheres, que durante sua vida fértil tiveram seus filhos de acordo com as Taxas Específicas de Fecundidade por Idade do período em estudo, e não estiveram expostas a riscos de mortalidade desde o nascimento até o término do período fértil.

Taxa de mortalidade infantil

É definida como o número de óbitos de menores de um ano de idade (por mil nascidos vivos), em determinada área geográfica e período, e interpreta-se como a estimativa do risco de um nascido vivo morrer durante o seu primeiro ano de vida.

Taxa específica de fecundidade por idade

É geralmente calculada por grupo qüinqüenal de idade, desde os 15 até os 49 anos. A taxa resulta da divisão do número de filhos nascidos vivos de mulheres do grupo de idade, em um período de tempo próximo à data do Censo Demográfico, usualmente os últimos 12 meses, pelo total de mulheres do mesmo grupo etário.

Taxa média geométrica de incremento anual da população

É a taxa de crescimento da população, dada pela expressão:

$$i = \sqrt[n]{\frac{P_{(t+n)}}{P_{(t)}} - 1}$$

onde:

$P_{(t)}$ é a população no ano t, e

n é o tempo, medido em anos.

Tempo ininterrupto de residência

Número de anos completos, contados até 31 de julho de 2000, pelos quais a pessoa morava, sem interrupção, no município e Unidade da Federação de residência atual.

Tipo de curso mais elevado freqüentado

Para a pessoa que, na data de referência, não freqüentava, mas anteriormente havia freqüentado escola, foi pesquisado o curso mais elevado no qual concluiu pelo menos uma série, da seguinte forma:

- Alfabetização de adultos – curso destinado à alfabetização de jovens e adultos;
- Antigo Primário – para curso elementar;
- Antigo Ginásio – para curso médio 1º ciclo;
- Antigo Clássico, Científico, etc. – para curso médio 2º ciclo;
- Ensino Fundamental ou 1º Grau;
- Ensino Médio ou 2º Grau;

- Superior – Graduação;

- Mestrado ou Doutorado – curso de mestrado ou doutorado. Este item somente foi considerado se houvesse a posse do título de mestre, doutor ou a aprovação da tese, ainda que o diploma não tivesse sido expedido; e

- Nenhum – para a pessoa que:

• freqüentou mas não concluiu a 1ª série do ensino fundamental, 1º grau ou equivalente;
• freqüentou ou concluiu somente classe de alfabetização;
• freqüentou mas não concluiu curso de alfabetização de adultos; ou
• freqüentou somente creche.

Tipo de Curso que freqüentava

Foi pesquisado o curso que a pessoa freqüentava, na data de referência, podendo ser;

- Creche – estabelecimento, juridicamente regulamentado ou não, destinado a dar assistência diurna às crianças nas primeiras idades;

- Pré-escolar – curso destinado, geralmente, a crianças de idade inferior a 7 (sete) anos, recebendo denominações que variam de acordo com a região e o nível alcançado pelas crianças: maternal, jardim de infância, jardim I, etc.;

- Classe de alfabetização – curso destinado à alfabetização de crianças;

- Alfabetização de adultos – curso destinado à alfabetização de jovens e adultos;

- Ensino fundamental ou 1º grau – regular seriado – curso de ensino fundamental ou 1º grau, organizado em séries anuais;

- Ensino fundamental ou 1º grau – regular não-seriado – curso de ensino fundamental ou 1º grau, não-organizado em séries anuais, ou seja, organizado em regime de créditos, períodos letivos, semestres, fases, módulos, ciclos, etc.;

- Supletivo de ensino fundamental ou 1º grau – curso supletivo de ensino fundamental ou de 1º grau, seriado ou não;

- Ensino médio ou 2º grau – regular seriado - curso de ensino médio ou 2º grau, organizado em séries anuais;

- Ensino médio ou 2º grau – regular não-seriado – curso de ensino médio ou 2º grau, não-organizado em séries anuais, ou seja, organizado em regime de créditos, períodos letivos, semestres, fases, módulos, ciclos, etc.;

- Supletivo de ensino médio ou 2º grau – curso supletivo de ensino médio ou 2º grau, seriado ou não;

- Pré-vestibular – curso preparatório para prestar exames de ingresso em curso de graduação superior (exame vestibular); não foi considerado nessa situação a pessoa que freqüentava, simultaneamente, o pré-vestibular e o ensino médio do 2º grau;

- Superior - Graduação – curso de graduação universitária ou superior;

- Mestrado ou Doutorado – considerou-se também nessa situação a pessoa que estava matriculada para a preparação de tese.

Para a pessoa que estava freqüentando mais de um curso, considerou-se o mais elevado que freqüentava. Especificamente, a pessoa que freqüentava, simultaneamente, curso de ensino médio ou 2º grau e pré-vestibular foi considerada como matriculada no ensino médio ou 2º grau.

Trabalho

Considerou-se como trabalho em atividade econômica o exercício de:

- ocupação remunerada em dinheiro, produtos, mercadorias ou somente em benefícios – moradia, alimentação, roupas, etc. – na produção de bens ou serviços;

- ocupação remunerada em dinheiro ou benefícios – moradia, alimentação, roupas, etc. no serviço doméstico;

- ocupação sem remuneração na produção de bens e serviços, desenvolvida durante pelo menos uma hora na semana:

a) em ajuda a membro do domicílio que tinha trabalho como conta-própria ou empregador;

b) em ajuda a membro do domicílio que tinha trabalho como empregado na produção de bens primários (que compreende as atividades de agricultura, silvicultura, pecuária, extração vegetal ou mineral, caça, pesca e piscicultura); ou

c) como aprendiz ou estagiário;

- ocupação desenvolvida, durante pelo menos uma hora na semana, na produção de bens do ramo que compreende as atividades de agricultura, silvicultura, pecuária, extração vegetal, pesca e piscicultura, destinados à alimentação de, pelo menos, um morador do domicílio.

Trabalho principal

Considerou-se como principal o único trabalho que a pessoa tinha na semana de referência. Para a pessoa que tinha mais de um trabalho, ou seja, que estava ocupada em mais de um empreendimento nessa semana, adotaram-se os seguintes critérios, na ordem descrita, para definir o principal:

- o que a pessoa dedicava maior número de horas normalmente trabalhadas por semana, independentemente de ser remunerado ou não;

- o que a pessoa tinha há mais tempo, independentemente de ser remunerado ou não, no caso de igualdade no número de horas normalmente trabalhadas por semana; e

- o que proporcionava o maior rendimento, no caso de igualdade também no tempo de exercício dos trabalhos.

No caso de troca definitiva de trabalho durante a semana de referência, considerou-se como principal, o último trabalho exercido.

União conjugal

A união conjugal foi investigada, na data de referência, identificando se a pessoa:

- vivia em união – somente no caso em que o cônjuge ou companheiro (a) fosse morador (a) do domicílio;

- não vivia, mas já viveu em união – foram incluídos também os casos em que o cônjuge ou companheiro (a) perdeu a condição de morador (a); e

- nunca viveu em união.

6.4 Classificações

Quanto às classificações utilizadas para a aplicação de códigos aos quesitos investigados em aberto no Censo Demográfico 2000, ocorreu um amplo trabalho de revisão, tendo em vista objetivos de harmonização/padronização no plano nacional e internacional.

Os temas que passaram pela aplicação de códigos foram: Religião, Migração, Curso, Ocupação e Atividade. O tema Migração congregava cinco quesitos: "Qual é a unidade da federação ou país estrangeiro de nascimento?", "Qual é a unidade da federação ou país estrangeiro de residência anterior?", "Em que município residia em 31 de julho de 1995?", "Em que unidade da federação ou país estrangeiro residia em 31 de julho de 1995?", e "Em que município e unidade da federação, ou país estrangeiro, trabalha ou estuda?"

6.4.1 Curso

O quesito 4.35 do Censo Demográfico 2000 – Questionário da Amostra – que se refere à espécie de curso mais elevado concluído pelo entrevistado (Graduação, Mestrado ou Doutorado), teve seus códigos aplicados segundo duas classificações distintas de área de formação dos cursos de educação superior.

Os códigos foram aplicados tendo como base a classificação de áreas de formação usada nos censos anteriores, derivada da estrutura e do conteúdo da classificação adotada pelo Ministério da Educação – MEC – em levantamentos nas instituições de educação superior do País. Os resultados apresentados nos códigos de dois dígitos desta classificação estão disponíveis para estudos comparativos com os censos demográficos anteriores.

Numa segunda etapa, através de uma tabela de conversão, foram aplicados outros códigos ao quesito, tomando por base a Classificação de Áreas de Formação e Treinamento adotada pelo MEC no Censo da Educação Superior de 2000. Trata-se de classificação derivada da versão 1997, da Classificação Internacional Padronizada da Educação da UNESCO (*International Standard Classification of Education* – ISCED), e do detalhamento de áreas de educação e treinamento (códigos de três dígitos, correspondendo ao terceiro nível da classificação) proposto no documento conjunto da UNESCO, OCDE e EUROSTAT, denominado *Fields of education and training: manual* (ANDERSON; OLSSON,1999), com alguns ajustes para sua aplicação no sistema educacional brasileiro.

A classificação de Áreas de Formação dos Cursos de Educação Superior usada no Censo Demográfico 2000, em nível de três dígitos, segue a estrutura e o detalhamento da classificação do MEC. Tanto sua extensão quanto os respectivos conteúdos contemplam as mudanças estruturais que os campos profissionais e do conhecimento têm experimentado nos últimos anos, estando organizada com a seguinte estrutura hierárquica:

Quadro 6.1 - Estrutura da classificação de áreas de formação dos cursos de educação superior - 2000

Nome das áreas	Nível	Número de grupamentos	Identificação
Áreas gerais (Grandes áreas)	1º	8	1 dígito
Áreas específicas (Áreas)	2º	22	2 dígitos
Áreas detalhadas (Subáreas)	3º	90	3 dígitos

Para a aplicação de códigos no 3º nível definidos a três dígitos, o IBGE introduziu alguns ajustes de denominação, baseando-se na lista de cursos do MEC e também na associação das informações contidas no banco descritor da base de dados do Censo Demográfico 2000.

Os códigos aplicados ao quesito 4.35, em nível de três dígitos da nova classificação, possibilita a comparação dos resultados do Censo 2000 com informações do MEC e com outras fontes internacionais. Além disso, por ser mais detalhada, permite melhor compreensão dos estudos relacionados ao perfil educacional da sociedade brasileira, bem como os referentes ao mercado de trabalho.

A adoção pelo IBGE da mesma classificação usada pelo MEC, em levantamentos do sistema educacional, e a opção de tomar como referência a classificação internacional da UNESCO inserem-se no movimento de padronização das classificações usadas no sistema estatístico e nos registros e cadastros da Administração Pública, visando à articulação entre as diversas fontes nacionais e de compatibilização com as classificações internacionais, possibilitando comparações e agregações para atender a finalidades específicas. A Classificação de Áreas de Formação e Treinamento dos Cursos de Educação Superior do MEC foi analisada e referendada por um grupo de trabalho organizado pela Comissão Nacional de Classificação – CONCLA, do qual participaram representantes do Instituto Nacional de Estudos e Pesquisas Educacionais – INEP/ MEC, do IBGE, do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, da Coordenação do Aperfeiçoamento de Pessoal de Nível Superior – CAPES e da Financiadora de Estudos e Projetos – FINEP.

Cabe ressaltar que a classificação que vinha sendo adotada nos censos anteriores, em nível de dois dígitos, não é compatível com a atual classificação feita com três dígitos.

No CD-ROM que acompanha a metodologia, são apresentadas as estruturas completas das duas classificações utilizadas na aplicação de códigos ao quesito 4.35 do Questionário da Amostra do Censo 2000, ou sejam:

- Tabela de Códigos e Denominação dos Cursos de Educação Superior – nível de 2 dígitos; e

- Classificação de Áreas de Formação e Treinamento dos Cursos de Educação Superior 2000 – nível de 3 dígitos.

6.4.2 Religião

O IBGE e o Instituto Superior de Estudos da Religião – ISER – em parceria, desenvolveram a classificação de religião dos Censos Demográficos 1991 e 2000.

Como a aplicação de códigos do quesito religião se fez com base na declaração do entrevistado a um quesito aberto e o desenvolvimento tecnológico permitiu maior qualidade no nível de detalhamento das informações, o Censo Demográfico 2000 apresenta uma série de aperfeiçoamentos significativos na sua grade classificatória. São as seguintes as principais modificações:

- não foram consideradas as divisões agregadas externas ou estranhas ao campo pesquisado;

- foi utilizada como critério geral de nomeação dos agregados e subagregados uma aproximação com o vocabulário corrente dos declarantes;

- foram desagregadas tradições religiosas com genealogias e/ou características morfológicas diferenciadas;

- foram incluídas declarações de religião com freqüência e/ou visibilidade social relevante; e

- foi preservada a estrutura geral da classificação para manter a comparação em nível mais agregado com os censos demográficos anteriores e possibilitar a análise das séries históricas.

Desta forma, ao tomar como critério maior a preservação da memória da declaração recebida, o censo demográfico ajusta-se à tendência das padronizações nacional e internacional.

A classificação adotada viabiliza o desenvolvimento de estudos e a comparação com outras pesquisas em universos religiosos situados, regionais e nacionais. Possibilita ainda a comparação com informações de órgãos nacionais de estatística e de institutos de pesquisa de outros países sobre religiosidades tradicionais e globalizadas.

As classificações utilizadas para aplicação de códigos ao quesito que investigou a religião das pessoas, nos Censos de 1991 e 2000, estão apresentadas no anexo do CD-ROM.

6.4.3 Migração

A aplicação de códigos aos quesitos pesquisados em aberto no Questionário da Amostra, relativos ao tema Migração – 4.21, 4.23, 4.25, 4.26 e 4.27 – levou em conta, para o nível município, a estrutura de códigos elaborada no IBGE. Tendo em conta que a aplicação de códigos foi feita, simultaneamente, para unidade da federação e país estrangeiro – quesitos 4.21, 4.23, 4.26 e 4.27 - não foi possível utilizar a estrutura de códigos normalmente utilizada no IBGE, para identificação da unidade da federação. Os códigos efetivamente utilizados estão apresentados no anexo de CD-ROM.

Mais tarde, tendo como objetivo atender às recomendações da Divisão de Estatísticas da ONU, foram associados às declarações, através de uma tabela de conversão, os códigos dados pela *Standard Country or Areas Codes for Statistical Use* (1999).

6.4.4 Ocupação

O quesito 4.45 do Questionário da Amostra investigou a ocupação do trabalhador, ou seja, a tarefa ou função que a pessoa desenvolvia na data de referência do censo. Tecnicamente, a ocupação representa a agregação de empregos ou situações de trabalho similares quanto às atividades realizadas, pois o que existe são as atividades exercidas pelo indivíduo em um emprego ou outro tipo de relação de trabalho.

No censo demográfico, o indivíduo declara sua ocupação, que deverá ser objeto de aplicação de código para permitir o tratamento estatístico dessa informação.

A Classificação Brasileira de Ocupações – CBO –, de responsabilidade do Ministério do Trabalho e Emprego – MTE – é o documento que reconhece, nomeia e codifica os títulos e descreve as características das ocupações do mercado de trabalho brasileiro. Esta classificação tem como referência a *Clasificación Internacional Uniforme de Ocupaciones*: *CIUO-88* (1991), da OIT, o Cadastro Brasileiro de Ocupações, organizado pelo MTE em 1971, e a *Classificação da Mão-de-Obra do Setor Primário*, elaborada pelo Instituto de Planejamento Econômico e Social – IPEA, Instituto de Planejamento – IPLAN, Centro Nacional de Recursos Humanos – CNRH, em 1977, para o uso nos registros administrativos.

No Brasil, as informações provenientes de registros administrativos relativas às ocupações eram codificadas seguindo a estrutura da CBO. Entretanto, os dados censitários e de pesquisas domiciliares seguiam uma nomenclatura própria do IBGE, sem descrições. Esta multiplicidade de classificações dificultava a comparabilidade entre as diferentes fontes de informações produzidas no país, além de dificultar a comparação com estatísticas de outros países.

Desde 1996, assumindo que a classificação é uma das bases do sistema estatístico, o Ministério do Trabalho e Emprego e o IBGE iniciaram o processo de revisão da CBO, conduzido no âmbito da CONCLA, com o objetivo de unificar as classificações nacionais e aproximá-las ao máximo da CIUO–88, permitindo, desta forma, comparar com maior eficácia as informações relacionadas ao mercado de trabalho nacional e internacional.

O Censo Demográfico 2000 utilizou pela primeira vez a Classificação Brasileira de Ocupações – CBO, adaptada para as pesquisas domiciliares[4], daqui por diante denominada CBO Domiciliar. Como esta classificação representou uma grande mudança em relação àquela utilizada nos censos anteriores, as ocupações tiveram os códigos aplicados segundo a classificação utilizada no Censo Demográfico 1991, para permitir a avaliação das diferenças existentes.

[4] No momento de realização do Censo Demográfico 2000, o desenvolvimento da nova CBO ainda não havia terminado, mas estava em etapa de validação de estrutura e descrição das famílias ocupacionais. Portanto, o IBGE utilizou a estrutura definida em outubro de 1999, que posteriormente sofreu pequenas alterações até a oficialização da *Classificação Brasileira de Ocupações: CBO-2002*, pelo MTE, em setembro de 2002.

Tanto a estrutura de códigos da CBO Domiciliar como a utilizada no Censo de 1991, estão apresentadas no anexo de CD-ROM, que acompanha o volume da metodologia.

a) Classificação utilizada no Censo de 1991

O IBGE adotou nos censos demográficos, até 1991, uma classificação de ocupações própria, que tinha como referência a Classificação de Programas dos Censos da América – COTA –, desenvolvida na década de 1960. Essa classificação de ocupações, revista por ocasião dos censos, ou seja, a cada dez anos, não era facilmente comparável com a CBO e com as classificações adotadas internacionalmente, inclusive com a CIUO. Durante o intervalo censitário, armazenava-se grande quantidade de declarações de ocupações que iam paulatinamente sendo analisadas, em termos quantitativos e de formato, convertendo-se assim na fonte geradora de um arquivo sistematizado de declarações de ocupação (descritores), que permitia incorporar ou redimensionar a classificação, acompanhando o desenvolvimento da força de trabalho e satisfazendo as necessidades do usuário de dados.

b) Classificação Brasileira de Ocupações Adaptada para as Pesquisas Domiciliares – CBO Domiciliar

A CBO descreve e ordena as ocupações dentro de uma estrutura hierarquizada que permite agregar as informações referentes à força de trabalho segundo características que dizem respeito às funções, tarefas e obrigações do trabalhador e ao conteúdo de seu trabalho (conhecimentos, habilidades e outros requisitos exigidos para o exercício da ocupação). Esta classificação toma como referência a *Clasificación Internacional Uniforme de Ocupaciones: CIUO–88* (1991). Ainda que apoiada nas orientações da CIUO, procurou-se garantir a representação de características ocupacionais do mercado de trabalho brasileiro em suas múltiplas manifestações, resultando em uma correspondência entre a nova Classificação de Ocupações e a classificação de referência, a CIUO–88 , a dois dígitos.

A unidade de observação é o emprego ou posto de trabalho, dentro de um conjunto de empregos mais amplo (campo profissional), onde o trabalhador terá mais facilidade em se movimentar. Estes conjuntos de empregos são identificados por processos, funções ou ramos de atividades.

A unidade de classificação mais desagregada é o grupo de base ou família ocupacional, identificado e definido em função da natureza do trabalho executado.

A estrutura da CBO Domiciliar baseia-se em quatro níveis, como pode-se verificar no quadro abaixo, sendo semelhante à CIUO-88:

Quadro 6.2 - Estrutura da Classificação Brasileira de Ocupações - CBO domiciliar - 2000

Estrutura	CBO domiciliar	CBO 2002	CIUO/88
Grandes grupos (GG) – 1 dígito	10	10	10
Subgrupos principais (SGP) – 2 dígitos	48	47	28
Subgrupos (SG) - 3 dígitos	176	192	116
Grupos de Base ou Famílias (GB) - 4 dígitos	510	596	390

A CBO Domiciliar mantém-se idêntica à CBO no nível mais agregado - grande grupo. Algumas famílias ocupacionais, subgrupos e subgrupos principais foram reagrupados tendo em vista as dificuldades de serem captados com precisão nas pesquisas domiciliares.

Os dez grandes grupos (GG) formam o nível mais agregado da classificação. Cada um apresenta um conjunto de famílias ocupacionais agregadas por nível de competência[5] e similaridade nas atividades executadas, reunindo amplas áreas de trabalho, mais do que tipos específicos. Cada grande grupo é identificado através do primeiro dígito do código.

O Subgrupo Principal (SGP) é um agrupamento mais restrito que o grande grupo, num total de 48 títulos. Reúne ocupações que apresentam estreito parentesco, tanto em relação à natureza do trabalho quanto aos níveis de qualificação exigidos. Fogem a este princípio os grupos residuais que, por sua natureza, agregam ocupações de parentescos mais fluidos.

O subgrupo SG, identificado pelo 3º dígito, indica, de forma ampla, o domínio dos campos profissionais de famílias ocupacionais agregadas. Para esta agregação levou-se em consideração, além dos critérios básicos, as necessidades de sistematização das informações, análises, previsões e programas, assim como outros aspectos que estabelecem afinidades baseadas na atividade econômica, na especificação profissional, no nível de escolaridade, nos tipos de equipamentos utilizados, na situação de complementaridade e em outros fatores, num total de 176 subgrupos.

O Grupo de Base ou Família Ocupacional (GB) é a unidade do sistema de classificação para efeitos práticos, reunindo um conjunto de ocupações similares, totalizando 510 famílias. Define-se Ocupação como o conjunto de postos de trabalho substancialmente iguais quanto à natureza e às qualificações exigidas. Constitui tarefas, obrigações e responsabilidades atribuídas a cada trabalhador.

No anexo de CD-ROM estão definidos cada um dos dez Grandes Grupos, sendo apresentados, também, seus principais componentes.

6.4.5 Atividade

O quesito 4.46 do questionário da amostra investigou a atividade econômica em que a pessoa exercia seu trabalho principal. Como categoria conceitual, a atividade econômica refere-se à caracterização das unidades de produção em suas ações de combinação de recursos para a geração de bens e serviços. Assim, a informação solicitada neste quesito é a atividade principal da unidade de produção (empresa, negócio próprio, instituição ou entidade governamental ou privada) em que a pessoa trabalhou na semana de referência. O tratamento estatístico desta informação requer, como etapa preliminar, sua associação a categorias definidas em uma classificação de atividades econômicas, o que foi feito através da aplicação de códigos, procedimento que consiste em atribuir à descrição dada pelo informante o código da categoria a que corresponde na classificação.

---

[5] O conceito de **competência** tem duas dimensões: o **nível de competência,** que é função da complexidade, amplitude e responsabilidade das atividades desenvolvidas no trabalho, e o **domínio (ou especialização) da competência**, relacionado a características como área de conhecimento, função, atividade econômica, processo produtivo, equipamentos, bens produzidos, que identificarão o tipo de profissão ou ocupação.

O quesito sobre a atividade econômica foi codificado simultaneamente em duas classificações: a Classificação Nacional de Atividades Econômicas, adaptada a pesquisas domiciliares – CNAE Domiciliar, e a Classificação de Atividades do Censo de 1991. O procedimento de dupla codificação teve por objetivo garantir que a transição para uma nova base classificatória, com estrutura bastante diferenciada da anterior, pudesse ser feita preservando a comparação com a série histórica.

O quesito sobre a atividade econômica é o principal elemento de ligação entre as estatísticas domiciliares e as estatísticas levantadas na ótica das unidades de produção, o que reforça a importância do uso de um sistema de classificação padronizado. Por outro lado, as pesquisas domiciliares, e o censo demográfico em particular, são as fontes mais abrangentes de informações sobre o mercado de trabalho do País, em cuja caracterização é fundamental a identificação da atividade econômica.

a) classificação de atividades do Censo de 1991

A tabela de códigos de atividades, usada no Censo de 1991, detalha 169 classes de atividades econômicas agrupadas em 25 categorias de tabulação; a cada classe é atribuído um código de três dígitos. A organização desta tabela, preparada para a codificação do quesito de atividade econômica dos censos anteriores, embora apoiada nas classificações, então usadas nas pesquisas econômicas do IBGE e em versões anteriores da *Clasificación Industrial Internacional Uniforme de todas las Actividades Económicas – CIUU*, não mantém uma correspondência exata, em qualquer nível, com estas classificações, o que dificultava tanto a articulação entre as estatísticas de pesquisas econômicas e domiciliares, como as comparações internacionais. O uso da mesma tabela e dos mesmos descritores no Censo de 1991 e nas pesquisas domiciliares, por outro lado, foi importante fator de continuidade e comparação entre as estatísticas destas fontes. A relação dos códigos usados no Censo de 1991 é apresentada no anexo de CD-ROM.

b) classificação Nacional de Atividades Econômicas – CNAE Domiciliar

O Censo de 2000 adotou uma nova classificação na aplicação de códigos às declarações ao quesito 4.46 do Questionário da Amostra, a Classificação Nacional de Atividades Econômicas – CNAE Domiciliar, dentro dos seguintes objetivos:

- atualização e melhoria de qualidade do instrumento de classificação;

- ampliação da articulação com outras fontes no tocante à informação sobre atividade econômica; e

- adequação às comparações internacionais.

c) a CNAE e a padronização nacional

A CNAE é a classificação de referência para a produção de estatísticas sobre a atividade econômica. Com esta classificação, publicada no Diário Oficial da União em dezembro de 1994, teve início a padronização dos códigos de atividades econômicas usados no País, tanto no sistema estatístico como em cadastros e registros da Administração Pública, acompanhada de harmonização com a classificação internacional. A padronização de classificações e conceitos é condição necessária para a articulação das informações no tempo (séries temporais) e entre fontes nacionais e internacionais.

A CNAE é derivada da *Clasificación Industrial International Uniforme de todas las Actividades Econômicas – CIIU*, revisão 3 (1990). Sob a gestão do Departamento de Estatística das Nações Unidas – UNDP – esta classificação tem seu desenvolvimento centrado no objetivo de servir como instrumento para a harmonização na produção e disseminação de estatísticas em nível internacional. A decisão de adotar a CIIU como referência refletiu a preocupação com a comparabilidade internacional das estatísticas nacionais.

Tal como a CIIU, a CNAE está estruturada em quatro níveis hierárquicos. Nos dois primeiros níveis, denominados seção e divisão, a CNAE segue a estrutura da CIIU, inclusive mantendo os mesmos códigos e denominações. Nos níveis seguintes, grupo e classe, a CNAE introduz maior detalhamento, levando em conta a estrutura do País. A CNAE define 218 grupos e 564 classes, enquanto na CIIU 3, os quantitativos são de 159 e 292, respectivamente.

d) adaptação da CNAE para pesquisas domiciliares

O detalhamento das 564 classes da CNAE mostrou-se, em muitos casos, inadequado para pesquisas domiciliares. No Brasil, o censo demográfico e demais pesquisas domiciliares são realizados por entrevistas, e a pergunta sobre a atividade econômica resume-se à descrição dada pelo informante. Estas descrições, via de regra, não detalham, em toda extensão, as características da atividade, que são necessárias para a identificação de determinadas classes da CNAE, como é o caso, por exemplo, das atividades industriais. Constatou-se, portanto, a impossibilidade de adoção da CNAE em todo o detalhamento das classes. Por outro lado, o nível de grupo (três dígitos) revelou-se ora com agregação excessiva para o que é importante captar numa pesquisa domiciliar, ora com excesso de detalhamento. Optou-se, então, por um detalhamento próprio para pesquisas domiciliares, reagrupando ou abrindo classes da CNAE. A esta estrutura, que é idêntica à CNAE nos dois primeiros níveis – seção e divisão, denominou-se CNAE Domiciliar.

e) estrutura da CNAE Domiciliar

Os critérios usados no desenho da CNAE Domiciliar foram:

1- manter a mesma estrutura da CNAE nos dois primeiros níveis: seção e divisão, inclusive os códigos, com exceção das divisões 51 – Comércio Atacadista e 52 – Comércio Varejista e Reparação de Objetos Pessoais e Domésticos, que foram consolidadas em uma única divisão: 53 – Comércio por Atacado e a Varejo e Reparação de Objetos Pessoais e Domésticos. Considerou-se, no caso do comércio, que a informação prestada nas pesquisas domiciliares dificilmente faria menção ou incorporaria corretamente os conceitos de atacado e varejo da CNAE;

2- abaixo do nível de divisão, dois dígitos, abrir classes específicas para as pesquisas domiciliares, reagrupando ou detalhando classes da CNAE, conforme a necessidade, dentro dos seguintes parâmetros:

   - relevância da atividade;

   - descrição usualmente dada por informantes de pesquisas domiciliares com os elementos necessários para a correta codificação com o grau de detalhe proposto; e

- atividades com as pesquisas domiciliares como fonte única ou principal de informação, como no caso de comércio ambulante.

A definição das classes da CNAE Domiciliar sofreu adaptações, em função da necessidade de ajuste com a CAES Mercosur, como será explicado adiante.

Para o Censo de 2000, foram ainda definidas novas classes, em função de questões que surgiram no processo de codificação das respostas do quesito 4.46. Foi encontrado grande número de casos em que a descrição dada pelo informante sobre a atividade não tinha todos os elementos para sua associação a uma determinada classe, mas era clara sua vinculação a um grupo de classes compreendidas numa mesma divisão. Optou-se, nestes casos, por criar nova classe com especificação mais genérica dentro da divisão, para evitar uma perda maior de informação, caso fosse codificada como atividade não especificada. Com este procedimento, foram criadas sete novas classes na estrutura da CNAE Domiciliar, listadas a seguir:

- 18999 – Confecção de artigos do vestuário e acessórios, sob medida ou não;

- 45999 – Atividades de construção - especialidade não definida;

- 53999 – Atividades de comércio não especificadas;

- 55999 – Atividades de alojamento e/ou alimentação não definidas;

- 60999 – Transporte terrestre - modalidade não especificada;

- 75999 – Administração pública, esfera de governo não especificada;

- 80999 – Atividades de ensino não especificadas; e

- 85999 – Atividades de saúde não especificadas.

É importante notar que estas classes foram criadas para não se perder a informação referenciada à divisão (dois dígitos), mas introduzem um viés em relação às classes originalmente definidas como compondo as divisões na CNAE Domiciliar.

O quadro a seguir resume a estrutura da CNAE Domiciliar usada no Censo 2000.

Quadro 6.3 - Estrutura da Classificação Nacional de Atividades Econômicas - CNAE - domiciliar - 2000

Nome	Nível	Número de grupamentos	Identificação
Seção	1º	17	Código alfabético de 1 dígito
Divisão	2º	58	Código numérico de 2 dígitos
Classe	3º	214	Código numérico de 5 dígitos

A estrutura detalhada da CNAE Domiciliar – códigos e denominações – é apresentada no anexo de CD-ROM, que acompanha a metodologia.

f) CAES Mercosur

A metodologia da CNAE Domiciliar foi discutida no âmbito do Projeto Censo Mercosul, resultando na definição da *Clasificación de Actividades Económicas para Encuestas Sociodemográficas del Mercosur – CAES Mercosur*, articulada com a *Clasificación Industrial Internacional Uniforme de todas las Actividades Econômicas – CIIU*, revisão 3 (1990), dentro dos princípios usados na CNAE Domiciliar. Todos os países do Grupo já usam a CIIU como referência no desenvolvimento de suas classificações nacionais ou para uso direto da classificação.

A CAES Mercosur reconhece os três níveis de classificação descritos anteriormente para a CNAE Domiciliar: seção, também conhecido como categoria de tabulação, em número de 17, idêntico à CIIU 3; divisão, composto por dois dígitos numéricos, correspondendo às mesmas categorias da CIIU, exceto pela junção do comércio atacadista e varejista; e classe, composto por dois dígitos adicionais, correspondendo à maior desagregação da classificação. Trata de nível de abertura próprio da CAES Mercosur, que guarda relação com aberturas de grupos, classes ou sua combinação na CIIU 3.

Os países participantes do Projeto Censo Mercosul acordaram em usar a CAES Mercosur como ferramenta classificatória nas pesquisas domiciliares com quesito sobre a atividade econômica, e/ou fornecer as informações sobre atividades econômicas referenciadas a esta classificação.

O Brasil ajustou as aberturas das classes da CNAE Domiciliar, de forma que passassem a corresponder às desagregações de classes da CAES Mercosur. A CNAE Domiciliar tem, portanto, as características de uma classificação derivada da CAES Mercosur, com maior detalhe no nível das classes.

A seguir, quadro comparativo das estruturas da CAES Mercosur, CIIU 3 e CNAE Domiciliar.

Quadro 6.4 - Comparação entre a Clasificación Industrial Internacional Uniforme - CIIU 3 -, Classificación de Actividades Económicas para Encuestas Sociodemográficas del Mercosur - CAES Mercosur - e Classificação Nacional de Atividades Econômicas - CNAE domiciliar - 2000

	CIIU 3	CAES Mercosur	CNAE domiciliar
Categoria de tabulação (Seção)	17	17	
Divisão	60	59	58
Grupo	159	-	-
Classe	292	125	214

A tabela com a estrutura detalhada da CAES Mercosur e sua correspondência com a CIIU 3 e a CNAE Domiciliar é apresentada no anexo de CD-ROM.

g) formação de códigos da CNAE Domiciliar

Nível seção ou categoria de tabulação

Código alfabético de uma letra, idêntico ao da CNAE, CAES Mercosur e CIIU3.

Nível divisão

Códigos de dois dígitos numéricos, idênticos aos da CNAE, CAES e CIIU 3, a menos de:

- a divisão 53 da CNAE Domiciliar e CAES Mercosur, que é uma junção das divisões 51 e 52 da CNAE e CIIU 3; e

- a divisão 12 – Extração de urânio e tório da CIIU 3 e CAES Mercosur, que na CNAE é tratada agregada à divisão 13.

Nível classe

Como regra geral: códigos de cinco dígitos com a seguinte composição:

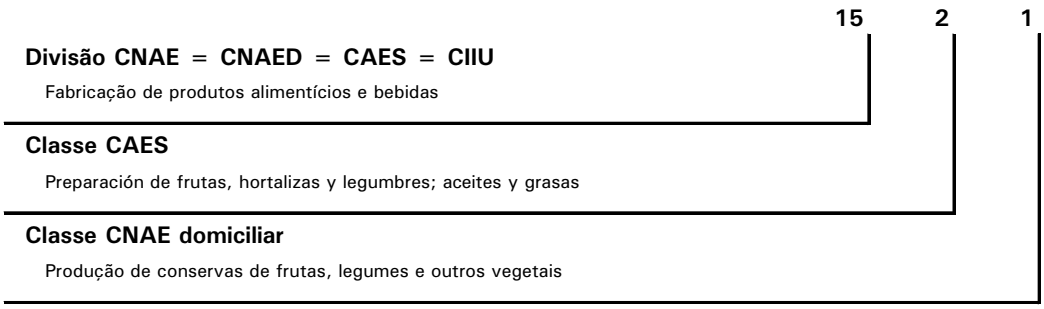
- dois dígitos da divisão;

- dois dígitos da CAES Mercosur, com a seguinte regra de formação: 00, quando não detalha a divisão; 0 + 1 a 9, de forma seqüencial, nos casos de detalhamento; e

- um dígito da CNAE Domiciliar, com a seguinte regra de formação: 0, quando não detalha a classe CAES, e 1 a 9, de forma seqüencial, nos casos de detalhamento

Exemplo de formação de código CNAE Domiciliar, refletindo a estrutura hierárquica e a articulação com CNAE, CAES e CIIU:

Quadro 6.5 - Formação do Código da Classificação Nacional de Atividades Econômicas - CNAE domiciliar - 2000


15
2
1
Divisão CNAE = CNAED = CAES = CIIU
Fabricação de produtos alimentícios e bebidas
Classe CAES
Preparación de frutas, hortalizas y legumbres; aceites y grasas
Classe CNAE domiciliar
Produção de conservas de frutas, legumes e outros vegetais

Exceções à regra geral

- em função de a classe CAES 0101 Agricultura ter sido aberta em 18 classes na CNAE Domiciliar, as classes CNAE Domiciliar da divisão 01 – Agricultura, pecuária, silvicultura e exploração florestal – têm a seguinte correspondência com a CAES:

CAES CNAE Domiciliar

0101 01101 a 01118

0102 01201 a 01209

0103 01300

0104 01401 e 01402

0105 01500

– nas classes de descrição genérica criadas para apuração do Censo 2000, os dois dígitos da divisão são seguidos por 999.

h) uso da CNAE Domiciliar

A CNAE Domiciliar foi usada pela primeira vez no Censo Demográfico 2000. A nova classificação representa uma mudança bastante ampla em relação às classificações usadas anteriormente nos censos e demais pesquisas demográficas. A inserção das classificações usadas no censo e demais pesquisas domiciliares, no movimento de padronização nacional e de harmonização internacional, tem como fator positivo a ampliação das possibilidades de articulação entre as estatísticas derivadas destas fontes e as provenientes de outras fontes, inclusive internacionais. A contrapartida a este avanço qualitativo são as restrições na comparabilidade das séries históricas, dada a impossibilidade de se estabelecer uma correspondência exata entre as categorias mais desagregadas da CNAE Domiciliar e as da classificação de atividades usada no Censo Demográfico 1991 e nas pesquisas domiciliares da década de 1990.

Para evitar solução de continuidade com a série dos censos demográficos passados, optou-se por codificar o quesito 4.46 segundo as duas classificações: a do Censo 1991 e a CNAE Domiciliar, possibilitando a leitura dos resultados do Censo 2000 nas duas classificações.

i) aplicação de códigos e banco de descritores

A aplicação da classificação de atividades econômicas é feita através da atribuição, a cada descrição, do código da categoria a que corresponde na classificação usada, através de procedimentos de codificação. Um instrumento fundamental na realização desta tarefa é o banco de descritores. Trata-se de uma lista a mais exaustiva possível das atividades contidas em cada classe da classificação, inclusive suas possíveis formas de descrição. O banco de descritores tem a dupla função de apoiar a codificação e garantir sua qualidade, inclusive pela homogeneidade de critérios e interpretações.

Tendo em vista a decisão de codificar o quesito 4.46, tanto na CNAE-Domiciliar como na Classificação de Atividades do Censo de 1991, foi preparado um banco de descritores referenciado simultaneamente aos códigos mais detalhados das duas classificações. Assim, para cada descrição associou-se um código de cinco dígitos da CNAE Domiciliar e outro de três dígitos da Classificação do Censo de 1991. O banco de descritores original do Censo 2000 foi construído a partir dos descritores usados no Censo de 1991 e nas pesquisas domiciliares da década de 1990, acrescentando-se a esta relação descrições de atividades provenientes das pesquisas econômicas.

Referências

XI RECENSEAMENTO geral do Brasil: manual de delimitação dos setores de 2000. Rio de Janeiro: IBGE, Diretoria de Geociências, 1998.

ANDERSON, R.; OLSSON, A. *Fields of education and training*: manual. Stockholm, Statistics Sweden, 1999.

ARAÚJO, T. C. N. S. *A classificação de "cor" nas pesquisas do IBGE*: notas para discussão. *Cadernos de Pesquisa*, São Paulo, n. 63, p. 14-16, nov. 1987.

BRASIL. Decreto nº 82.263, de 13 de setembro de 1978. Dá nova denominação aos atuais parques nacionais do Xingu e de Tumucumaque. *Diário Oficial [da] República Federativa do Brasil,* Poder Executivo, Brasília, DF, 14 set. 1978. p. 14933. Col. 1.

_____. Decreto nº 84.337, de 21 de dezembro de 1979. Cria a reserva indígena de Parabubure, no Estado do Mato Grosso, e dá outras providências. *Diário Oficial [da] República Federativa do Brasil,* Poder Executivo, Brasília, DF, 21 dez. 1979. p. 19584. Col. 1. Retificada no Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF, 27 dez. 1979. p. 19900. Col. 2.

_____. Lei nº 6.902, de 27 de abril de 1981. Dispõe sobre a criação de estações ecológicas, a proteção ambiental e dá outras providências. *Diário Oficial [da] República Federativa do Brasil,* Poder Executivo, Brasília, DF, 28 abr. 1981. p. 7557. Col. 1.

_____. Lei nº 6.938, de 31 de agosto de 1981. Dispõe sobre a política nacional do meio ambiente, seus fins e mecanismo de formulação e aplicação e dá outras providências. *Diário Oficial [da] República Federativa do Brasil,* Poder Executivo, Brasília, DF, 2 set. 1981. p. 16509. Col. 1.

_____. Decreto nº 89.336, de 31 de janeiro de 1984. Dispõe sobre as reservas ecológicas e áreas de relevante interesse ecológico, e dá outras previdências. *Diário Oficial [da] República Federativa do Brasil,* Poder Executivo, Brasília, DF, 1 fev. 1984. p. 1572. Col. 1.

_____. Decreto nº 94.946, de 23 de setembro de 1987. Regulamenta o item 1 do artigo 17 da Lei nº 6.001, de 19 de dezembro de 1973. *Diário Oficial [da] República Federativa do Brasil,* Poder Executivo, Brasília, DF, 24 set. 1987. p. 15594. Col. 1.

_____. Lei nº 7.853, de 24 de outubro de 1989. Dispõe sobre o apoio às pessoas portadoras de deficiência, sua integração social, sobre a Coordenadoria Nacional para Integração da Pessoa Portadora de Deficiência (CORDE), institui a tutela jurisdicional de interesses coletivos ou difusos dessas pessoas, disciplina a atuação do Ministério Público, define crimes, e dá outras providencias. *Diário Oficial [da] República Federativa do Brasil,* Poder Executivo, Brasília, DF, 25 out. 1989. Seção I.

_____. Lei complementar nº 94, de 19 de fevereiro de 1998. Autoriza o Poder Executivo a criar a Região Integrada de Desenvolvimento do Distrito Federal e Entorno – RIDE e instruir o programa especial de desenvolvimento do entorno do Distrito Federal, e dá outras providências. *Diário Oficial [da] República Federativa do Brasil,* Poder Executivo, Brasília, DF, 20 fev. 1998. p. 1. Col. 1.

BRASS, W. et al. *The demography of tropical Africa*. Princeton: Princeton University Press, 1968.

CASTIGLIONI, A.; DUCHENE, J. *Une mesure indirecte des migrations internationales*: l´influence du scheme de fecondité sur la structure par âge des émigrés. Louvain la Neuve, Belgique: CIACO, 1986. (Working paper, n. 133).

CENSO DEMOGRÁFICO 2000: características da população e dos domicílios: resultados do universo. Rio de Janeiro: IBGE, 2001. Acompanha CD-ROM.

CENSO DEMOGRÁFICO 2000: características da população e dos domicílios: resultados do universo. Rio de Janeiro: IBGE, 2001. Acompanha CD-ROM.

_____: instruções para revisão do trabalho da base operacional geográfica – TR 23. Rio de Janeiro: IBGE, 2000. 32 p.

_____: manual do recenseador – CD 1.09. Rio de Janeiro: IBGE, 2000. 151 p.

_____: primeiros resultados da amostra. Rio de Janeiro: IBGE, 2002. 2 pt.

CHACKIEL, Juan. Los temas demográficos en los censos de América Latina. In: SEMINÁRIO SOBRE O CENSO 2000 NO MERCOSUL, 2., 1997, Rio de Janeiro, RJ. *O censo 2000 no MERCOSUL*: segundo seminário de trabalho. Rio de Janeiro: IBGE, 1998. p. 107-118.

CLASIFICACION de actividades económicas para encuestas sociodemográficas del Mercosur – CAES-Mercosur. Disponível em: <http://www.ibge.gov.br/concla/cnaedomiciliar/cnaedom.shtm>. Acesso em: 1 dez. 2003.

CLASIFICACION industrial internacional uniforme de todas las actividades económicas - CIIU. Nueva York: Naciones Unidas, Oficina de Estadística, 1990. 212 p. (Informes estadísticos. Serie M, n. 4, rev. 3).

CLASIFICACION internacional uniforme de ocupaciones: CIUO-88. Ginebra: Oficina Internacional del Trabajo, 1991. 523 p.

CLASSIFICAÇÃO brasileira de ocupações: CBO-2002. Brasília, DF: Ministério do Trabalho e Emprego, Secretaria de Políticas Públicas de Emprego, 2002. 3 v.

CLASSIFICAÇÃO brasileira de ocupações: CBO-domiciliar [Censo 2000]. Disponível em: <http://www.ibge.gov.br/concla/ocupacao/cbo/cbo.shtm>. Acesso em: ago. 2001.

CLASSIFICAÇÃO da mão-de-obra do setor primário. Brasília, DF: IPEA, Instituto de Planejamento, Centro Nacional de Recursos Humanos, 1977. (Estudos para o planejamento, n. 17).

CLASSIFICAÇÃO nacional de atividades econômicas – CNAE. Rio de Janeiro: IBGE, 2002. 344 p.

CLASSIFICAÇÃO nacional de atividades econômicas – CNAE-domiciliar. Rio de Janeiro: IBGE, 2002. Disponível em: <http://www.ibge.gov.br/concla/cnaedomiciliar/cnaedom.shtm>. Acesso em: ago. 2001.

COSTA, V. G. Em busca da redefinição de favelas e assemelhados: o projeto caracterização dos setores de baixa renda. In: ENCONTRO NACIONAL DE GEÓGRAFOS, 13., 2002, João Pessoa. *Anais...* [São Paulo]: AGB, 2002. 1 CD-ROM.

DIVISÃO do Brasil em mesorregiões e microrregiões geográficas. Rio de Janeiro: IBGE, 1990. v.1: Brasil.

DIVISÃO do Brasil em micro-regiões homogêneas 1968. Rio de Janeiro: IBGE, 1970. 564 p.

DIVISÃO territorial do Brasil 2001. Rio de Janeiro: IBGE, Diretoria de Geociências, 2002. Trabalho inédito.

FAIJER, D. J. Algunas notas sobre la estimación de la migración internacional. *Notas de Población*, Santiago de Chile, ano 15, n. 43, p. 25-46, abr. 1987.

GUIDELINES and principles for the development of disability statistics. New York: United Nations, 2001. (Series Y, n. 10).

HILL, K. Estimación de la emigración por edades a partir de la información sobre residencia de hermanos. *Notas de Población*, Santiago de Chile, ano 7, n. 21, p. 71-89, dic. 1979.

IBGE. Resolução do Presidente – PR nº 11/90, de 5 de junho de 1990. Altera os códigos geográficos de meso e microrregiões de que trata o Quadro III, anexo à R. PR-51/89 e atualiza a divisão territorial até a presente data. *Boletim de Serviço [do] IBGE*, Rio de Janeiro, v. 38, p. 1, jun. 1990.

INTEGRANDO conceitos, classificações e experiências: o projeto do Censo de 2000 no Mercosul. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1999. 10 p.

INTERNATIONAL classification of functioning, disability and health – IFC. Geneva: World Health Organization, 2001.

INTERNATIONAL standard classification of education - ISCED. [Paris]: UNESCO, 1997. Disponível em: <http://www.uis.unesco.org/en/act/act_p/isced.html>. Acesso em: 5 ago. 2003.

INTERNATIONAL standard classification of occupations - ISCO-88. Geneva: International Labour Organization, [198?]. Disponível em: <http://www.ilo.org/public/english/bureau/stat/class/isco.htm>. Acesso em: 5 ago. 2003.

MBOGONI, M.; SYNNEBORN, M. *General disability measures in developing countries:* relationship to purposes for measurement. 2003. Trabalho apresentado no encontro do Washington Group on Disability Statistics, Ottawa, 2003.

OLIVEIRA, J. de C. O retorno da migração internacional como componente da dinâmica demográfica brasileira – até quando? In: MIGRAÇÕES internacionais: contribuições para políticas. Brasília: Comissão Nacional de População e Desenvolvimento, 2001. Trabalho apresentado no Seminário sobre Migrações Internacionais - Contribuições para Políticas, 2000, Brasília.

PEREIRA, N. O. M. Experiência brasileira na investigação dos grupos étnicos e populações afrodescendentes nas pesquisas domiciliares, 1872/2000. In: ENCUENTRO INTERNACIONAL TODOS CONTAMOS: LOS CENSOS Y LA INCLUSIÓN SOCIAL, 2., 2002, Lima. *Anais...* Lima: Instituto Nacional de Estadísticas e Informática, 2003. 1 CD-ROM.

_____. A religião nos censos brasileiros: informações preliminares do Censo Demográfico 2000. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA PARA O PROGRESSO DA CIÊNCIA, 54., 2002, Goiânia, GO. *Anais...* [Rio de Janeiro]: SBPC, 2002. Trabalho apresentado no Simpósio O País passa por Mutação Religiosa, 2002. 1 CD-ROM.

PEREIRA, N. O. M. Experiência brasileira na investigação dos grupos étnicos e populações afrodescendentes nas pesquisas domiciliares, 1872/2000. In: ENCUENTRO INTERNACIONAL TODOS CONTAMOS: LOS CENSOS Y LA INCLUSIÓN SOCIAL, 2., 2002, Lima. *Anais...* Lima: Instituto Nacional de Estadísticas e Informática, 2003. 1 CD-ROM.

_____; FLORIDO, A. C.; FERNANDES, M. Z. População residente em terras indígenas: características básicas censitárias, 1991/2000. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 13., 2002, Ouro Preto. *Anais...* [São Paulo]: ABEP, 2002. Tema: Violências, o Estado e a qualidade de vida da população brasileira. 1 CD-ROM.

PRIMEIRA prova piloto do censo demográfico 2000: relatório sobre a primeira prova piloto do censo demográfico 2000. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1997.

PRINCIPLES and recommendations for population and housing censuses. New York: United Nations, 1998. (Statistical papers. Series M, n. 67/ rev. 1).

SANDSTRÖM, A.; WRETMAN, J. H.; WALDÉN, B. Variance estimators of Gini coefficient-probability sampling. *Journal of Business & Economic Statistics*, Washington, D.C., v. 6, n. 1, p. 113-119, 1988.

SCHWARTZMAN, S. *Cor, raça e origem no Brasil.* 1999. Trabalho inédito.

SEGUNDA prova piloto do censo demográfico 2000: relatório de avaliação da segunda prova piloto do censo demográfico 2000: 1ª parte. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1998. 61 p.

_____: relatório de avaliação da segunda prova piloto do censo demográfico 2000: 2ª parte. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1998. 18 p.

SINOPSE estatística da educação superior 2000. Brasília, DF: INEP, 2001. 400 p.

SINOPSE preliminar do censo demográfico 2000. Rio de Janeiro: IBGE, v. 7, 2000. Acompanha 1 CD-ROM.

SOMOZA, J. Una idea para estimar la población emigrante por sexo y edad en el censo de un País. *Notas de Población*, Santiago de Chile, ano 5, n. 15, p. 89-106, 1977.

STANDARD country or areas codes for statistics use. New York: United Nations, Statistics Division, 1999. (Statistics and statistical methods publications, Series M, n.49/rev. 4). Disponível em: <http://unstat.un.org/unsd/methods/m49/m49alpha.htm>. Acesso em: 13 fev. 2002.

TABULAÇÃO avançada do censo demográfico 2000: resultados preliminares da amostra. Rio de Janeiro: IBGE, 2002.

7 Base Operacional Geográfica

Retratar o perfil demográfico, social e econômico da população, em um país com as dimensões e diferenciações regionais do Brasil, exige um cuidadoso planejamento das operações da base operacional geográfica, que tem início com:

- a caracterização do território nacional, em seus aspectos fundamentais para o adequado dimensionamento do volume e custos da operação de coleta, e

- a delimitação e descrição de unidades mínimas de coleta, de forma a garantir o perfeito reconhecimento pelo Recenseador de sua área de trabalho, evitando omissões e/ou duplicidades que podem prejudicar a cobertura dos domicílios.

A promulgação da Constituição Federal de 1988 redefiniu o papel da União, dos Estados e dos Municípios, levando estes dois últimos a assumirem responsabilidades quanto aos planos de ordenamento e gestão do território. Inseriu-se neste contexto uma crescente demanda por informações em recortes territoriais mais desagregados, a fim de que as estatísticas atendam às exigências do planejamento local. Assim, a unidade mínima de coleta – setor censitário – passou a ser demandada não apenas para cumprir seu papel original de suporte à coleta, mas também como unidade de construção de áreas para o planejamento de ações, tanto do setor público quanto do privado.

Uma das grandes preocupações que se tem ao planejar um Censo é evitar problemas de cobertura, o que se obtém com o aperfeiçoamento dos mapas utilizados na operação e com um sistema de arquivos e cadastros que, juntos, constituem o que se convencionou denominar Base Operacional Geográfica. Os componentes da base são:

- Mapa Municipal (MM) – mapa de um município em escala que varia de 1:50 000 a 1:250 000, contendo a visão geral do município, mostrando detalhadamente sua área rural com a identificação dos limites de distritos e subdistritos, os principais acidentes físicos, localidades, fazendas etc.; identifica áreas urbanas, ou seja, as cidades e as vilas, como pequenas áreas sem definição de ruas, quadras etc.;
- Mapa Municipal Estatístico (MME) – mapa municipal ao qual foi adicionada a malha de setores censitários rurais, e incluídas informações sobre o perímetro urbano com a numeração dos respectivos setores;
- Mapa de Localidade (ML) – conjunto de plantas urbanas, em escalas que variam de 1:2 000 a 1:10 000, que compõe as áreas urbanas de um município, com a identificação das ruas, avenidas, e eventuais marcos de referência (prédios públicos, igrejas, clubes etc.), limites de bairros, subdistritos, regiões administrativas, perímetro urbano etc.;
- Mapa de Localidade Estatístico (MLE) – mapa de localidade ao qual foi adicionada a malha de setores censitários urbanos, com a numeração dos respectivos setores;
- Sistema BOG (Base Operacional Geográfica) – sistema para apoio aos processos de aquisição, tratamento, gerenciamento e atualização dos dados, na construção da base do Censo Demográfico; compreende os cadastros e arquivos relativos às divisões territoriais institucionalizadas (UF, município, distrito e subdistrito), às áreas de apuração (aglomerados subnormais, bairros, aglomerados rurais e áreas especiais), às áreas de coleta (setor censitário), além da descrição de setores e comparabilidade;
- Arquivo de setor – registra todas as informações necessárias à comparabilidade; caracterização das alterações sofridas pelos setores de 1996, indicando os critérios aplicados na formação dos setores de 2000; relação entre setores de 1996 e 2000; subordinação político-administrativa da área do setor nas duas épocas; e, finalmente, composição da malha setorial de 2000 com as possíveis amarrações às áreas de apuração; e
- Arquivo de descrição do perímetro do setor – descreve, em detalhe, as linhas definidoras dos limites do setor.

Os avanços tecnológicos, especialmente aqueles voltados para a aquisição, tratamento e divulgação da informação geográfica, promoveram a disseminação de ferramentas de sistemas de informações geográficas, para os quais as bases cartográficas e os dados referenciados a áreas geográficas constituem requisitos fundamentais para as análises, e subsídios à tomada de decisão na implementação de planos de desenvolvimento. Face a este cenário, a base operacional para o Censo de 2000 incorporou requisitos exigidos, não apenas para as operações de coleta do Censo e de construção/atualização de cadastros para a seleção das amostras das pesquisas domiciliares no período intercensitário, mas também para sua utilização pelas novas tecnologias existentes tanto no IBGE, para divulgação dos resultados, como na comunidade usuária.

7.1 Diretrizes

A Base Operacional Geográfica contemplou as diversas alterações ocorridas no quadro territorial brasileiro, quanto às alterações da divisão político-administrativa e aos novos arranjos sociais e econômicos. Assim, as operações vinculadas à produção das bases territoriais de 2000 consideraram, dentre outros aspectos:

- as divisões e subdivisões político-administrativas, de forma a garantir que um setor censitário estivesse contido em uma única unidade administrativa (município, distrito);

- as áreas com limites indefinidos e problemáticos, buscando-se definir e descrever com acuidade os limites dos setores censitários, prevenindo duplicação e/ou omissão de unidades;

- as áreas de rápido crescimento/esvaziamento populacional, onde os padrões de assentamento populacional estão mudando rapidamente;

- as áreas com população esparsa que constituem problemas para a delimitação de setores;

- a adequada identificação da situação urbana e rural, segundo a legislação vigente;

- a adequada avaliação do ambiente físico e de áreas especiais (reservas indígenas, áreas de conservação ambiental, dentre outras);

- o mapeamento de áreas sem urbanização adequada (como aglomerados subnormais);

- a busca por soluções de tratamento cartográfico para áreas onde ainda não se dispunha de mapeamento básico adequado;

- a construção do Cadastro de Segmentos de Logradouros por setor censitário, ferramenta de controle e avaliação da coleta, que teve por objetivo fornecer informação detalhada dos intervalos de endereços (segmentos) e respectivos Códigos de Endereçamento Postal (CEP), existentes em cada setor censitário; e

- a definição de um padrão de armazenamento das informações, em formato digital, mais condizente com as exigências da sociedade atual, e que permitisse adições de informações produzidas interna e externamente ao IBGE.

Embora estivesse claro para os responsáveis pela elaboração da base operacional que a completa informatização seria altamente benéfica ao processo, restrições de prazo e custo obrigaram a reduzir este aperfeiçoamento. Optou-se por priorizar a atualização e digitalização, e a conversão das informações gráficas para meio digital, da base territorial dos municípios com população acima de 50 000 habitantes, o que representou uma cobertura de mais de 60% da população brasileira.

Considerando que a informatização dos procedimentos tem prazos de implementação longos, adotou-se a descentralização de atividades no âmbito do IBGE, através das Divisões Estaduais de Pesquisas – DIPEQs – e Divisões de Geociências – DIGEOs – com o estabelecimento de parcerias com outros órgãos federais, estaduais e municipais, e com a terceirização de atividades, através da contratação de serviços e de pessoal. Como essas opções são afetadas pelo tipo de situação – urbana ou rural – os trabalhos de preparo da base operacional 2000 foram desenvolvidos em duas vertentes, a saber:

Vertente Urbana - Apoiada em mapeamento cadastral - escala de 1:2 000 a 1:10 000 - é produzido nas esferas estadual e municipal, por órgãos públicos e empresas privadas. A partir desses mapas foram elaborados os mapas urbanos e mapas urbanos estatísticos, onde se representam os logradouros e limites de setores, e que estão associados a cadastros que incluem os logradouros e descrições de setores urbanos, bairros e outros; e

Vertente Rural - Apoiada no mapeamento topográfico sistemático existente, que tem responsabilidade nacional dividida entre o IBGE e a Diretoria de Serviço Geográfico do Comando do Exército – DSG, foram elaborados os mapas municipais e os mapas municipais estatísticos, onde estão representados os elementos físicos naturais e artificiais do terreno, tais como rios, estradas, localidades, os limites municipais, distritais e de setores, que estão associados a cadastros de topônimos, de localidades, de propriedades rurais, de áreas especiais, de descrições de setores e outros.

A tabela 7.1, a seguir, apresenta a divisão político-administrativa vigente para a coleta do Censo de 2000 e permite avaliar o volume de trabalho envolvido na preparação da Base Operacional Geográfica.

Tabela 7.1 - Áreas, por tipo, segundo as Grandes Regiões e Unidades da Federação - 2000

Grandes Regiões e Unidades da Federação	Áreas, por tipo				
	Mesorregiões	Microrregiões	Divisão político-administrativa		
			Regiões Metropolitanas e RIDE (1)	Municípios	Distritos
Brasil	**136**	**547**	**21**	**5 507**	**9 516**
Norte	**20**	**64**	**1**	**449**	**586**
Rondônia	2	8	-	52	61
Acre	2	5	-	22	22
Amazonas	4	13	-	62	81
Roraima	2	4	-	15	15
Pará	6	22	1	143	227
Amapá	2	4	-	16	29
Tocantins	2	8	-	139	151
Nordeste	**42**	**188**	**5**	**1 787**	**2 988**
Maranhão	5	21	1	217	244
Piauí	4	15	-	221	221
Ceará	7	33	1	184	729
Rio Grande do Norte	4	19	1	166	186
Paraíba	4	23	-	223	278
Pernambuco	5	19	1	(2) 185	(2) 382
Alagoas	3	13	-	101	116
Sergipe	3	13	-	75	83
Bahia	7	32	1	415	749
Sudeste	**37**	**160**	**6**	**1 666**	**3 068**
Minas Gerais	12	66	2	853	1 535
Espírito Santo	4	13	1	77	237
Rio de Janeiro	6	18	1	91	276
São Paulo	15	63	2	645	**1 020**
Sul	**23**	**94**	**7**	**1 159**	**2 189**
Paraná	10	39	3	399	727
Santa Catarina	6	20	3	293	427
Rio Grande do Sul	7	35	1	467	1 035
Centro-Oeste	**14**	**41**	**2**	**446**	**685**
Mato Grosso do Sul	4	11	-	77	165
Mato Grosso	4	11	-	126	216
Goiás	5	18	1	242	303
Distrito Federal	1	1	1	1	1

Fonte: IBGE, Diretoria de Geociências, Departamento de Estruturas Territoriais.

(1) Região Integrada de Desenvolvimento do Distrito Federal e entorno. (2) Inclui o Distrito de Fernando de Noronha.

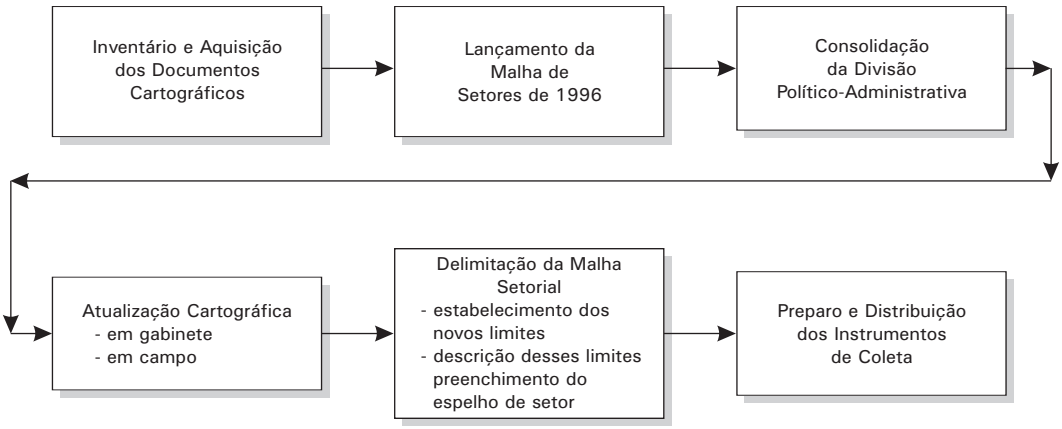
7.2 Elaboração da Base Operacional Geográfica

A elaboração da base operacional geográfica do Censo 2000 consistiu na delimitação dos setores, na elaboração dos cadastros, na produção dos instrumentos de suporte à coleta e na divulgação dos produtos.

Por motivos operacionais, esta delimitação teve por princípio básico a manutenção das áreas dos setores censitários de 1996 utilizados para a Contagem da População. Esta fase de trabalho teve início com a identificação dos limites dos setores nos mapas municipais e mapas urbanos e concluiu com o preenchimento dos formulários de Descrição do Perímetro do Setor e do Espelho de Setores.

O fluxo apresentado a seguir ilustra, de modo simplificado, a seqüência de elaboração da base operacional do Censo 2000. Descreve-se a seguir, em maior detalhe, cada uma destas atividades.

Figura 7.1 - Fluxo de elaboração da Base Operacional Geográfica


Inventário e Aquisição dos Documentos Cartográficos
Lançamento da Malha de Setores de 1996
Consolidação da Divisão Político-Administrativa
Atualização Cartográfica
- em gabinete
- em campo
Delimitação da Malha Setorial
- estabelecimento dos novos limites
- descrição desses limites preenchimento do espelho de setor
Preparo e Distribuição dos Instrumentos de Coleta

Inventário e aquisição dos documentos cartográficos

Consistiu na elaboração de um inventário das bases cartográficas, em escala topográfica e cadastral, que já se dispunha no IBGE, e na busca de documentação produzida externamente por órgãos oficiais (federais, estaduais e municipais) de cartografia e planejamento e mesmo na iniciativa privada, através de parcerias.

Lançamento da malha de setores de 1996

Consistiu no lançamento dos limites dos setores da Contagem da População 1996, no mapa mais recente disponível. No caso dos setores rurais, este lançamento foi realizado no mapa municipal e, para os setores urbanos, utilizou-se o mapa de localidade.

Consolidação da Divisão Político-Administrativa

Consistiu no reconhecimento e análise dos limites municipais e sua adequada representação em documento cartográfico.

Atualização cartográfica

A existência de originais das folhas topográficas não foi suficiente para resolver o problema de confecção da base municipal digital pois o mapeamento topográfico existente estava, em grande parte, defasado, acarretando problemas para o reconhecimento e caracterização do espaço geográfico representado. A desatualização desse mapeamento de referência, aliada ao pouco tempo disponível para a realização da atividade de atualização cartográfica no campo, fez com que os trabalhos ficassem restritos à finalidade censitária, atendendo somente ao levantamento de elementos de referência dos setores censitários, às principais localidades, à rede viária e outros elementos de importância relativa, que foram modificados no terreno.

A atualização cartográfica envolveu:

- uma avaliação do grau de desatualização do mapeamento municipal rural e do mapeamento municipal urbano, utilizando documentos cartográficos existentes, para verificação das divergências e/ou existência de elementos que poderiam ser confirmados em campo; esta atividade foi realizada em gabinete; e

- uma operação de campo, cujo objetivo foi o reconhecimento no terreno e o adequado posicionamento de elementos fundamentais para a efetiva identificação da unidade de coleta. Consistiu no levantamento dos acidentes geográficos existentes no terreno, utilizando-se de aparelhos receptores GPS (*Global Position System*) para o registro de coordenadas. Os elementos atualizados foram: vias de acesso, pontos de interseção, pontes, edificações, topônimos, elevações, rios, aglomerados humanos, logradouros etc.

As informações cartográficas levantadas em campo, na área rural, com o navegador GPS, foram registradas em uma Ficha de Atualização de Campo – FAC, onde constam as coordenadas geográficas, o topônimo do elemento cartográfico, sua posição relativa no terreno, além de outras informações. Estes registros compuseram o banco de dados de informações atualizadas, que é acessado pelo Sistema de Cartografia Semi-automatizada – SISCART para a produção do Mapa Municipal Digital – MMD atualizado. As atualizações foram incorporadas ao mapa municipal automaticamente.

Com o decorrer dos trabalhos, foram identificados pelas DIPEQs, alguns municípios com população abaixo de 50 000 habitantes que, por especificidades próprias, necessitaram também de atualizações em campo. A tabela abaixo demonstra o quantitativo de municípios que foram atualizados por regiões do país e por faixa de população.

Tabela 7.2 - Municípios, total e que passaram por atualização cartográfica, por classes de população do município, segundo as Grandes Regiões - 2000

Grandes Regiões	Municípios					
	Total	Que passaram por atualização cartográfica				
		Total atualizados	Percentual (%)	Classes de população do município		
				Municípios com população acima de 25 000 habitantes (coluna 4)	Municípios com população entre 10 000 e 25 000 habitantes (coluna 5)	Municípios com população abaixo de 10 000 habitantes (coluna 6)
Brasil	**5 507**	**1 304**	**23,67**	**995**	**202**	**107**
Norte	449	165	36,74	89	34	42
Nordeste	1 787	451	25,23	361	82	8
Sudeste	1 666	408	24,48	334	48	26
Sul	1 159	149	12,85	148	1	0
Centro-Oeste	446	131	29,37	63	37	31

Fonte: IBGE, Diretoria de Geociências, Departamento de Cartografia.

Embora a meta inicial fosse a atualização cartográfica dos municípios com mais de 50 000 habitantes, optou-se por abranger também aqueles acima de 25 000 (coluna 4). A agilização dos processos, contudo, permitiu um ganho extra com a atualização de municípios de menor contingente populacional, quantificados nas colunas 5 e 6.

Delimitação da malha setorial

O setor censitário de 2000, como já informado, teve por princípio básico a manutenção das áreas dos setores censitários de 1996, desde que esta situação não provocasse problemas na coleta, permitindo-se nesses casos a subdivisão ou agregação de setores de 1996, inteiros ou em parte, com base nos critérios estabelecidos e indicados a seguir:

- Setor de Área Urbana, Urbanizada – mínimo de 250 e máximo de 350 domicílios, de forma a propiciar a coleta do Censo Demográfico em até 30 dias;

- Setor de Área Urbana, Não-Urbanizada – mínimo de 150 e máximo de 250 domicílios ou, mínimo de 100 e máximo de 200 estabelecimentos agropecuários, com até 45 dias para coleta do Censo Demográfico;

- Setor de Área Rural – mínimo de 150 e máximo de 250 domicílios ou, mínimo de 100 e máximo de 200 estabelecimentos agropecuários ou área máxima de 500 km^2; para coleta do Censo Demográfico em até 45 dias e, em até 60 dias, para coleta do Censo Agropecuário;

- Setor Aglomerado Rural – mínimo de 51 e máximo de 200 domicílios;

- Setor de Especial Aglomerado Subnormal (favelas e similares) – mínimo de 51 e máximo de 350 domicílios;

- Setor Especial de Aldeia Indígena – mínimo de 20 habitantes indígenas em uma ou mais moradias;

- Setor Especial de Quartéis, Bases Militares etc. – mínimo de 50 moradores;

- Setor Especial de Alojamentos, Acampamentos etc. – mínimo de 50 moradores;

- Setor Especial de Penitenciárias, Colônias Penais, Presídios, Cadeias etc. – mínimo de 50 moradores; e

- Setor Especial de Asilos, Orfanatos, Conventos, Hospitais etc. – mínimo de 50 moradores.

Confrontando-se os limites dos setores de 1996 com os critérios quantitativos estabelecidos e com as informações disponíveis, em especial as obtidas em campo, as equipes estaduais procederam ao estabelecimento dos novos limites setoriais - a delimitação dos setores de 2000. A partir dos novos limites procedeu-se seu lançamento no mapeamento e ao preenchimento dos formulários para atualização dos cadastros.

Antes de realizar o preenchimento dos formulários, foi necessária a numeração dos setores. Para a numeração, a regra básica foi seqüencial crescente, dentro da menor área existente no código geográfico do IBGE, iniciando por 1. Desta forma, a atividade de numeração dos setores obedeceu aos seguintes passos:

• em um distrito, subdistrito, região administrativa ou zona, a numeração devia ser única, seqüencial crescente, iniciando por 1;

• iniciava-se a numeração no quadro urbano e continuava-se no quadro rural do mesmo distrito, subdistrito, região administrativa ou zona;

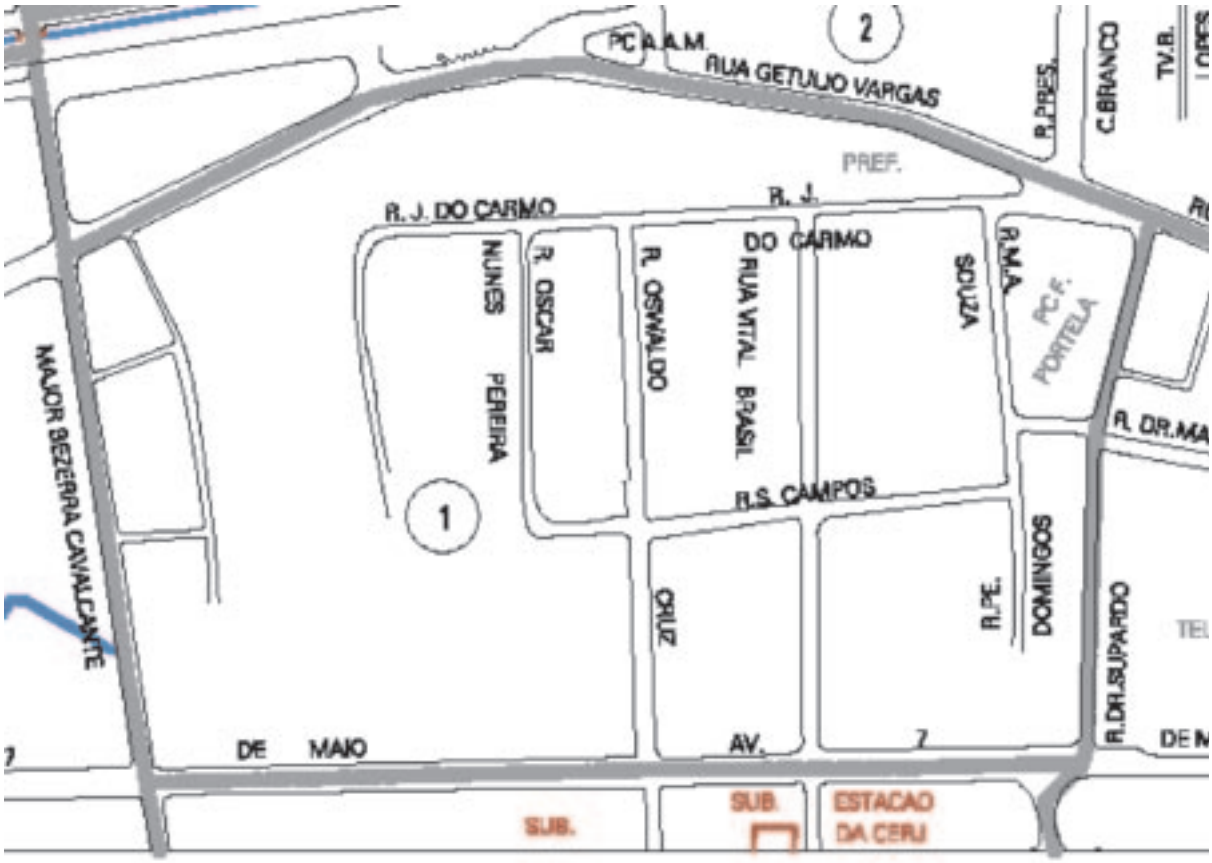
• iniciava-se a numeração partindo da área considerada como o centro da cidade ou da vila;

• numerava-se, a partir de 1, seguindo, de um modo geral, em espiral no sentido dos ponteiros do relógio e em direção à periferia; e

• após a numeração de todos os setores de um distrito, subdistrito, região administrativa ou zona, passava-se para o distrito, subdistrito, região administrativa ou zona seguinte, repetindo-se a numeração, novamente, a partir de 1, até que estivessem identificados todos os setores de todos os distritos, subdistritos, regiões administrativas ou zonas de um município.

A figura 7.2 ilustra a numeração de setores.

Figura 7.2 Resultado da numeração de setores

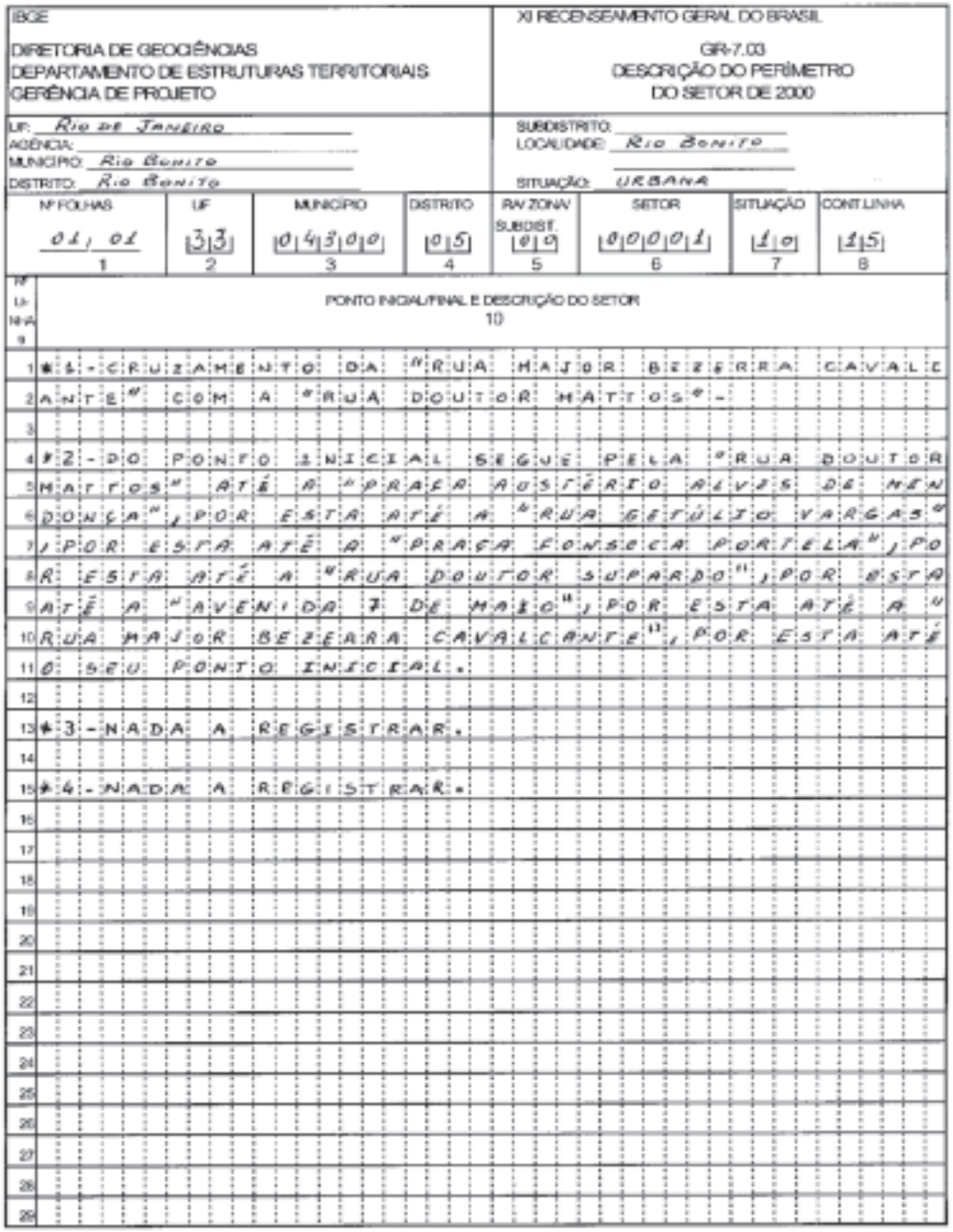


Procedeu-se, a seguir, à descrição dos limites de setores (perímetro do setor) levando-se em consideração:

- pontos de referência estáveis e de fácil identificação no terreno, como, por exemplo: logradouros, praças, linhas férreas, nas áreas urbanas; e divisores de águas, cumeadas de morros, estradas, sedes de fazendas, nas áreas rurais;
- respeito aos limites legais; e
- respeito aos limites de áreas de apuração (áreas especiais, aglomerados subnormais - favelas e similares – aglomerados rurais, aldeias indígenas, bairros e similares, áreas urbanizadas e áreas não urbanizadas).

A seguir, apresenta-se exemplo do formulário – Descrição do Perímetro do Setor de 2000 - GR 7.03.

Figura 7.3 Descrição do perímetro do setor



IBGE
DIRETORIA DE GEOCIÊNCIAS
DEPARTAMENTO DE ESTRUTURAS TERRITORIAIS
GERÊNCIA DE PROJETO

XI RECENSEAMENTO GERAL DO BRASIL
GR-7.03
DESCRIÇÃO DO PERÍMETRO DO SETOR DE 2000

UF: Rio de Janeiro
AGÊNCIA:
MUNICÍPIO: Rio Bonito
DISTRITO: Rio Bonito
SUBDISTRITO:
LOCALIDADE: Rio Bonito
SITUAÇÃO: URBANA

Nº FOLHAS	UF	MUNICÍPIO	DISTRITO	RA/ ZONA/ SUBDIST.	SETOR	SITUAÇÃO	CONT LINHA
01 / 01	33	04300	05	00	000041	10	15
1	2	3	4	5	6	7	8

Nº DE ORDEM

PONTO INICIAL/FINAL E DESCRIÇÃO DO SETOR
10

1 *1 - CRUZAMENTO DA "RUA MAJOR BEZERRA CAVALC
2 ANTE" COM A "RUA DOUTOR MATTOS" -
3
4 2 - DO PONTO INICIAL SEGUE PELA "RUA DOUTOR
5 MATTOS" ATÉ A "PRAÇA AUSTÉRIO ALVES DE MEN
6 DONÇA", POR ESTA ATÉ A "RUA GETÚLIO VARGAS"
7 , POR ESTA ATÉ A "PRAÇA FONSECA PORTELA", PO
8 R ESTA ATÉ A "RUA DOUTOR SUPARDO", POR ESTA
9 ATÉ A "AVENIDA 7 DE MAIO", POR ESTA ATÉ A "
10 RUA MAJOR BEZERRA CAVALCANTE", POR ESTA ATÉ
11 O SEU PONTO INICIAL.
12
13 *3 - NADA A REGISTRAR.
14
15 *4 - NADA A REGISTRAR.
16
17
18
19
20
21
22
23
24
25
26
27
28
29

Ao fim dos procedimentos de definição do limite do setor 2000, iniciou-se o preenchimento do formulário Espelho de Setores 2000 que, após digitado, deu origem ao arquivo conhecido como "Arquivo de Setores". O Espelho de Setores retrata para cada setor sua situação político administrativa, sua participação em áreas de apuração - áreas especiais (AESP), bairros (BAI), aglomerados subnormais (FAV) e aglomerados rurais (AGLOM) – e sua genealogia, ou seja, sua relação com os setores existentes em 1996. Desta forma, é por seu intermédio que se estabelecerá a comparabilidade entre os setores da Contagem da População 1996 e do Censo Demográfico 2000.

No preenchimento do formulário utilizaram-se os Códigos de Formação e de Situação e Tipo do Setor de 2000, descritos a seguir.

Código de formação do setor de 2000

É um código constituído por três dígitos com o objetivo de identificar as alterações sofridas pelo setor de 1996 para a formação dos setores de 2000, bem como a operação adotada, a saber:

- o primeiro dígito indica a composição da área do setor de 2000, em termos das áreas oriundas do setor de 1996, inteiras e/ou em partes;

- o segundo dígito indica o tipo de operação adotada para a formação do setor de 2000: manutenção, subdivisão ou agregação da área do setor de 1996. Além disso, contém código específico para identificar os setores de 1996 que foram extintos ou parcialmente alagados; e

- o terceiro dígito indica se houve modificação na subordinação político-administrativa ou na situação da área do setor de 1996, para a formação do setor de 2000.

O quadro 7.1 apresenta o conjunto de Códigos de Formação do Setor de 2000.

Quadro 7.1 - Códigos de formação do setor de 2000

Primeira posição		Segunda posição		Terceira posição			
Código	Composição da área	Código	Operação para a formação	Código	Subordinação		Situação de 2000
					De município de 2000	De distrito, subdistrito, RA ou zona	
1	Área do setor de 2000 constituída pela área do setor de 1996	1	Manutenção	1	Mantida a de 1996	Mantida a de 1996	Mantida a de 1996
2	Área do setor de 2000 constituída por parte do setor de 1996	2	Subdivisão por critério quantitativo de unidades de coleta	2	Mantida a de 1996	Mantida a de 1996	Mantida a de 1996
		3	Subdivisão por área superior a 500 km quadrados	3	Mantida a de 1996	Mantida a de 1996	Mantida a de 1996
		4	Subdivisão por limite de área de apuração	4	Mantida a de 1996	Mantida a de 1996	Mantida a de 1996
		5	Subdivisão por outro motivo	5	Mantida a de 1996	Mantida a de 1996	Mantida a de 1996
3	Área do setor de 2000 constituída pela junção de setores de 1996 inteiros	6	Agregação por critério quantitativo	6	Mantida a de 1996	Mantida a de 1996	Mantida a de 1996
4	Área do setor de 2000 constituída pela junção de partes de setores de 1996	7	Agregação por outro motivo				
5	Área do setor de 2000 constituída pela junção de área(s) de setor(es) de 1996 inteiro(s) com parte de setor(es) de 1996						
6	Área do setor de 2000 parcialmente alagada	8	Extinção de parte do setor de 1996 (apenas parte alagada)	0	-	-	-
0	Área do setor totalmente alagada	0	Extinção de todo o setor de 1996	0	-	-	-

Código de situação e tipo do setor de 2000

É o código constituído por dois dígitos que identifica a situação legal urbana ou rural do setor, bem como o seu tipo (especial ou não especial).

O Quadro 7.2 apresenta o conjunto de Códigos de Situação e Tipo do Setor de 2000.

Quadro 7.2 - Códigos de situação e tipo do setor de 2000

Especificação do tipo do setor	Códigos de situação do setor							
	Urbana			Rural				
	Cidade ou vila		Área urbana isolada	Aglomerado rural				Zona rural
	Área urbanizada	Área não urbanizada		De extensão urbana	Isolado			
					Povoado	Núcleo	Outros aglomerados	Exclusive aglomerado rural
Não especial	10	20	30	40	50	60	70	80
Especial de aglomerado subnormal (favelas e similares)	11	-	31	41	-	-	-	-
Especial de quartéis, bases militares, etc.	12	22	32	42	52	62	72	82
Especial de alojamentos, acampamentos, etc.	13	23	33	43	53	63	73	83
Especial de embarcações, barcos, navios, etc.	14	24	34	44	54	64	74	84
Especial de aldeia indígena	15	25	35	45	55	65	75	85
Especial de penitenciárias, colônias penais, presídios, etc.	16	26	36	46	56	66	76	86
Especial de asilos, orfanatos, conventos, hospitais, etc.	17	27	37	47	57	67	77	87

O preenchimento do formulário foi feito dentro do município de 2000, em ordem de Distrito, RA, Zona ou Subdistrito, conforme indicado:

UF: Registrado o código da Unidade da Federação que consta do relatório da Divisão Territorial do Brasil.

MUNICÍPIO: Registrado o código do município que consta do relatório da Divisão Territorial do Brasil.

DISTRITO (DIST): Registrado o código do Distrito a que pertence o setor de 2000, cuja informação foi retirada do relatório da Divisão Territorial do Brasil.

SUBDISTRITO (SDIST): Registrado o código da Região Administrativa, Subdistrito, Zona ou similar nos municípios onde existem estas subdivisões, de acordo com o relatório da Divisão Territorial do Brasil; nos demais municípios, registrou-se 00.

Campos de identificação do setor de 2000

NÚMERO DO SETOR (SETOR): Registrado o número do setor de 2000.

SITUAÇÃO E TIPO (SIT): Registrado o código da situação e tipo do setor de 2000.

ESTIMATIVA DE DOMICÍLIOS (EDOC): Registrado o total de domicílios estimados no setor de 2000. Quando o setor de 2000 foi formado por uma agregação ou por uma subdivisão de setores de 1996, este registro foi repetido em todas as linhas em que aparecer o setor de 1996.

ESTIMATIVA DE UNIDADES VISITADAS (EUV): Registrado o total de unidades visitadas no setor de 2000. Quando o setor de 2000 foi formado por uma agregação ou por uma subdivisão de setores de 1996, este registro foi repetido em todas as linhas em que aparecer o setor de 1996.

CÓDIGO DE FORMAÇÃO (FRM): Registrado o código de formação do setor de 2000.

CONTROLE DE FORMAÇÃO (CT): Registrado o total de setores de 1996 que participaram da formação do setor de 2000.

ÁREA ESPECIAL (AESP): Registrado o código de identificação da ÁREA ESPECIAL, retirado do relatório do Banco de Estruturas Territoriais.

BAIRRO OU SIMILAR (BAI): Registrado o código de identificação do BAIRRO ou SIMILAR, retirado do relatório do Banco de Estruturas Territoriais.

AGLOM. SUBNORMAL (FAV. E SIMILARES) (FAV): Registrado o código de identificação do AGLOMERADO SUBNORMAL, retirado do relatório do Banco de Estruturas Territoriais.

AGLOM. RURAL E ALDEIA INDÍGENA (AGLOM): Registrado o código de identificação do AGLOMERADO RURAL ou ALDEIA INDÍGENA, retirado do Relatório do Banco de Estruturas Territoriais.

Campos de identificação do setor de 1996

UF E MUNICÍPIO / SETOR DE 1996 (UF MUN): Registrado o código da Unidade da Federação e do Município, retirado do relatório do Banco de Estruturas Territoriais, ano-base de 1996.

CÓD. DIST., RA, SUBDIST. OU ZONA/SETOR DE 1996 (DIST. SDIST) Registrado o código do DISTRITO, RA, SUBDIST. ou ZONA, retirado do Relatório do Banco de Estruturas Territoriais, ano Base de 1996.

N. DO SETOR / SETOR DE 1996: Registrado o número do setor de 1996.

SITUAÇÃO E TIPO / SETOR DE 1996: Registrado o código de situação e tipo do setor de 1996.

DOM. E UNIDADES VISITADAS /SETOR DE 1996: As informações foram obtidas no ESPELHO DE SETORES 1996 ou através de consulta ao Banco de Dados da Base Operacional existentes nas Unidades Regionais.

A seguir, apresenta-se um exemplo de preenchimento do formulário Espelho de Setores 2000.

Quadro 7.3 - Preenchimento do formulário Espelho de Setores 2000

IBGE — Sistema de Delimitação de Setores

DGC — Espelho de Setores 2000 por Município

DETRE / GPRG

UF: 11 MUNICÍPIO: 98 - ESPIGÃO D´OESTE

DIST	SDIST	SETOR	SIT	EDOC	EUV	FRM	CT	AESP	BAI	FAV	AGLOM	UF	MUN	DIST	SDIST	SET	SIT
5	0	34	80	113	135	471	2					11	98	5	0	26	8
		34	80	113	135	471	2					11	98	5	0	27	8
		35	80	148	185	471	2					11	98	5	0	12	8
		35	80	148	185	471	2					11	98	5	0	27	8
		36	80	8	9	111	1					11	98	5	0	24	8
		37	80	0	0	111	1	7 340				11	98	5	0	33	8
		38	80	111	121	111	1					11	98	5	0	28	8
		39	80	61	88	111	1					11	98	5	0	34	8
		40	50	124	167	111	1				3	11	98	5	0	35	5
		41	80	119	144	111	1					11	98	5	0	29	8
		42	80	63	73	471	2					11	98	5	0	20	8
		42	80	63	73	471	2					11	98	5	0	25	8
		43	80	81	96	471	4					11	98	5	0	20	8
		43	80	81	96	471	4					11	98	5	0	21	8
		43	80	81	96	471	4					11	98	5	0	23	8
		43	80	81	96	471	4					11	98	5	0	25	8
		44	80	9	9	111	1	7 501				11	98	5	0	36	8
		45	80	0	0	111	1	7 501				11	98	5	0	30	8
		46	80	16	16	111	1	7 501				11	98	5	0	31	8

Preparo e distribuição dos instrumentos de coleta

A caderneta do setor reúne, em um só instrumento, vários documentos necessários à realização da coleta das informações. Nela são anexados o mapa e a descrição do perímetro do setor de modo a orientar o percurso do Recenseador. A caderneta serve, também, para acondicionar tanto as Folhas de Coleta quanto as Folhas de Domicílio Coletivo. Em sua capa, dentre outras informações, apresenta os totais de domicílios e pessoas recenseadas, representando um resumo da coleta do setor.

Foram impressos os seguintes quantitativos:

157.415 mapas de setores urbanos;

58.396 mapas de setores rurais; e

215.811 descrições de perímetros dos setores (urbanos e rurais).

As descrições de perímetro do setor foram integralmente impressas nas Divisões Estaduais de Pesquisa. Já a impressão dos mapas, se realizou de forma descentralizada - nas DIPEQs para a vertente urbana e nas Divisões de Geociências, no caso da vertente rural.

Este material foi repassado às agências que efetivamente realizaram a tarefa de montagem da Caderneta do Setor, colando a descrição na capa e o mapa do setor na contracapa da caderneta. A seguir são apresentados mapas e descrições de setores do Município de Rio Bonito, Rio de Janeiro.

Figura 7.4 - Mapa do Setor 1, situação urbana, Município de Rio Bonito (RJ)

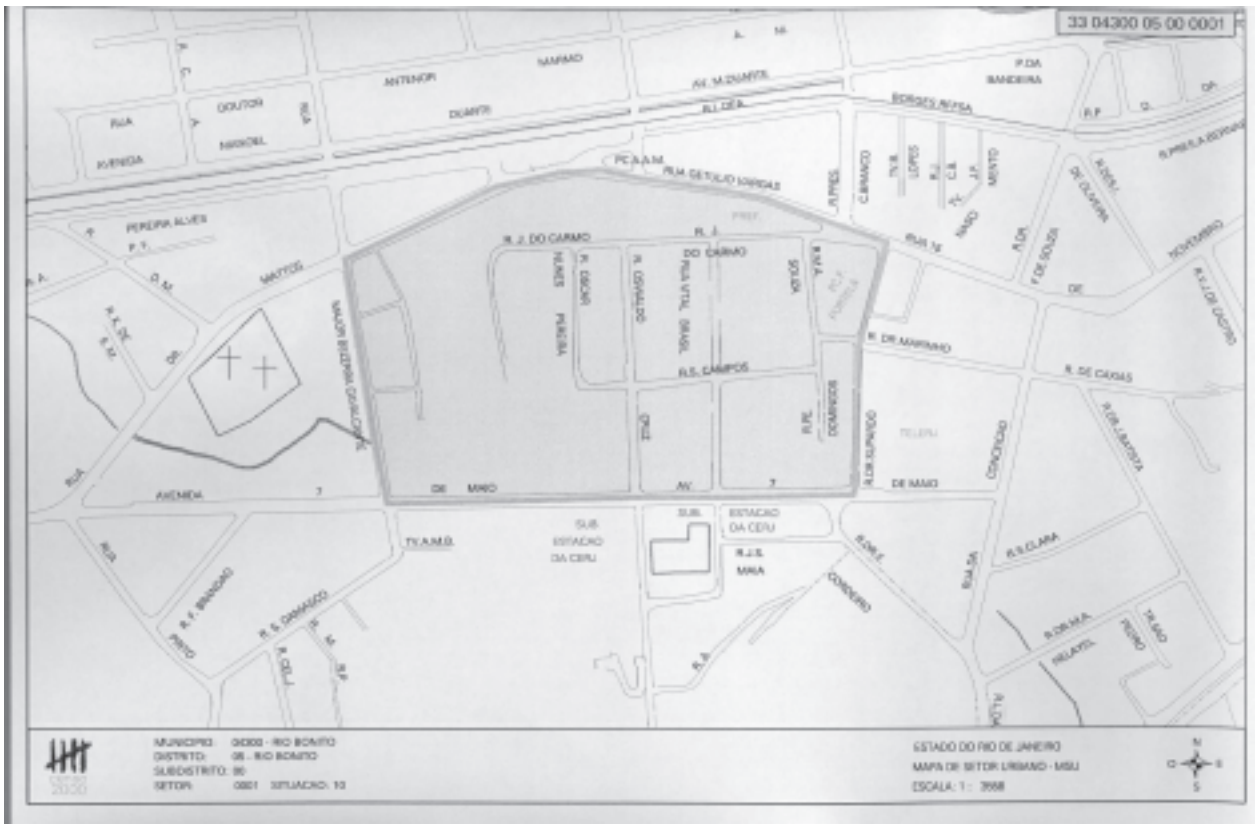


Figura 7.5- Descrição do Perímetro do Setor 1, situação urbana, Município de Rio Bonito (RJ)

CENSO 2000 Página: 1

UF:	Rio de Janeiro	33
MUNIC:	Rio Bonito	04380
DISTR.:	Rio Bonito	05
SDISTR.:		00
SITUAÇÃO:	10 - Urbana	SETOR: 0001

PONTO INICIAL E FINAL

CRUZAMENTO DA "RUA MAJOR BEZERRA CAVALCANTE" COM A "RUA DOUTOR MATTOS".

DESCRIÇÃO DO PERÍMETRO

DO PONTO INICIAL SEGUE PELA "RUA DOUTOR MATTOS" ATÉ A "PRAÇA ALISTÉRIO ALVES DE MENDONÇA", POR ESTA ATÉ A "RUA CETULIO VARGAS", POR ESTA ATÉ A "PRAÇA FONSECA PORTELA", POR ESTA ATÉ A "RUA DOUTOR SUPARDO", POR ESTA ATÉ A "AVENIDA 7 DE MAIO", POR ESTA ATÉ A "RUA MAJOR BEZERRA CAVALCANTE", POR ESTA ATÉ O SEU PONTO INICIAL.

SETORES A SEREM EXCLUÍDOS

NADA A REGISTRAR.

AGLOMERADOS RURAIS SOMENTE IDENTIFICADOS

NADA A REGISTRAR.

Figura 7.6 - Mapa do Setor 12, situação rural, Município de Rio Bonito (RJ)

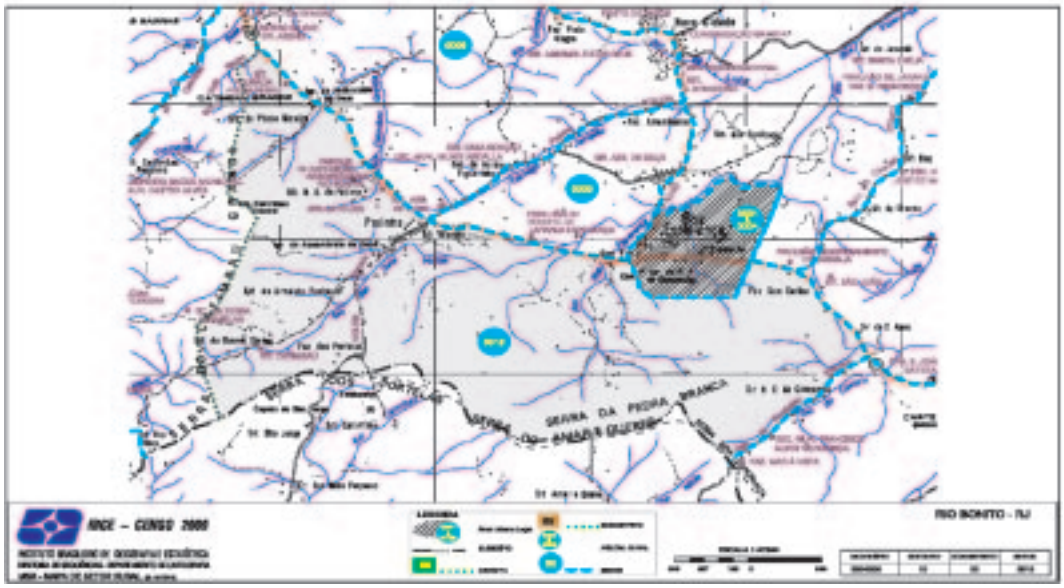
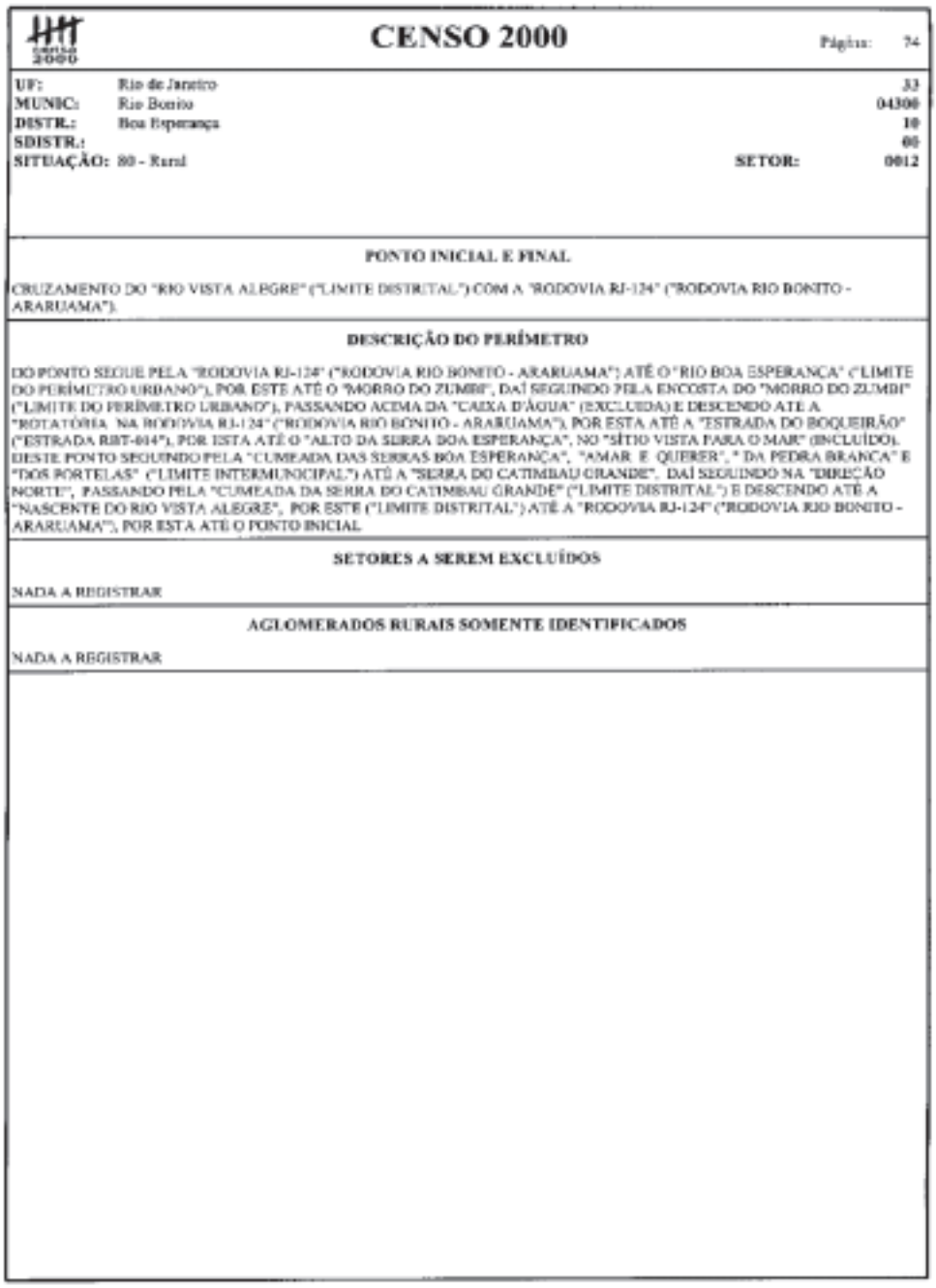


Figura 7.7 - Descrição do Perímetro do Setor 12, situação rural, Município de Rio Bonito (RJ)



CENSO 2000 Página: 74

UF: Rio de Janeiro 33
MUNIC: Rio Bonito 04300
DISTR.: Boa Esperança 10
SDISTR.: 00
SITUAÇÃO: 80 - Rural SETOR: 0012

PONTO INICIAL E FINAL

CRUZAMENTO DO "RIO VISTA ALEGRE" ("LIMITE DISTRITAL") COM A "RODOVIA RJ-124" ("RODOVIA RIO BONITO - ARARUAMA").

DESCRIÇÃO DO PERÍMETRO

DO PONTO SEGUE PELA "RODOVIA RJ-124" ("RODOVIA RIO BONITO - ARARUAMA") ATÉ O "RIO BOA ESPERANÇA" ("LIMITE DO PERÍMETRO URBANO"), POR ESTE ATÉ O "MORRO DO ZUMBI", DAÍ SEGUINDO PELA ENCOSTA DO "MORRO DO ZUMBI" ("LIMITE DO PERÍMETRO URBANO"), PASSANDO ACIMA DA "CAIXA D'ÁGUA" (EXCLUÍDA) E DESCENDO ATÉ A "ROTATÓRIA NA RODOVIA RJ-124" ("RODOVIA RIO BONITO - ARARUAMA"), POR ESTA ATÉ A "ESTRADA DO BOQUEIRÃO" ("ESTRADA RBT-014"), POR ESTA ATÉ O "ALTO DA SERRA BOA ESPERANÇA", NO "SÍTIO VISTA PARA O MAR" (INCLUÍDO). DESTE PONTO SEGUINDO PELA "CUMEADA DAS SERRAS BOA ESPERANÇA", "AMAR E QUERER", " DA PEDRA BRANCA" E "DOS PORTELAS" ("LIMITE INTERMUNICIPAL") ATÉ A "SERRA DO CATIMBAU GRANDE". DAÍ SEGUINDO NA "DIREÇÃO NORTE", PASSANDO PELA "CUMEADA DA SERRA DO CATIMBAU GRANDE" ("LIMITE DISTRITAL") E DESCENDO ATÉ A "NASCENTE DO RIO VISTA ALEGRE", POR ESTE ("LIMITE DISTRITAL") ATÉ A "RODOVIA RJ-124" ("RODOVIA RIO BONITO - ARARUAMA"), POR ESTA ATÉ O PONTO INICIAL

SETORES A SEREM EXCLUÍDOS

NADA A REGISTRAR

AGLOMERADOS RURAIS SOMENTE IDENTIFICADOS

NADA A REGISTRAR

Cadastro de Segmentos de Logradouros

Como parte dos trabalhos de criação da Base Operacional Geográfica do Censo 2000, foi elaborado um cadastro relacionando os trechos de logradouro identificados em cada setor censitário urbano. Denominado Cadastro de Segmentos de Logradouros, este arquivo constitui o embrião de um futuro Cadastro de Endereços para uso na coleta, não apenas do Censo, como também em todas as pesquisas do IBGE.

Embora não utilizado como ferramenta nas operações do Censo 2000, suas aplicações são inúmeras, como por exemplo a possibilidade de associação automática do endereço ao setor censitário, pré-requisito para a vinculação de informações sobre escolas, estabelecimentos comerciais e industriais e quaisquer outros identificados por um endereço, às informações socioeconômicas do Censo 2000.

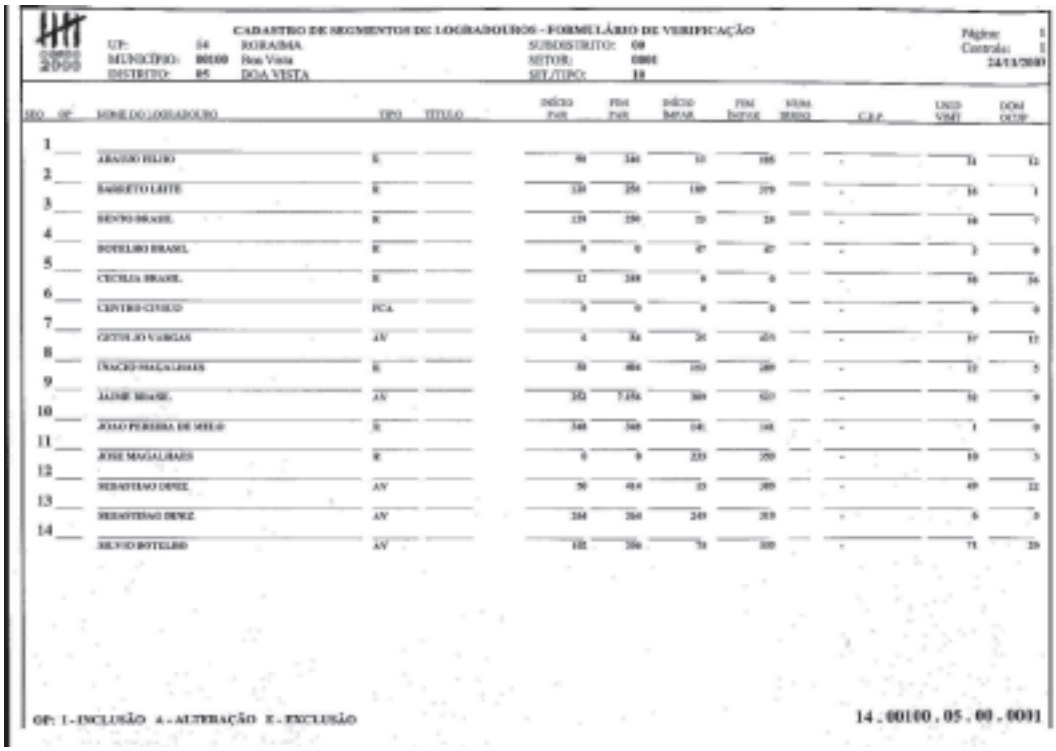
O projeto de construção do Cadastro teve início com a transcrição dos nomes dos logradouros e da numeração inicial e final contida em cada face de todas as quadras dos setores, registradas nas Folhas de Coleta da Contagem da População 1996, para formulário apropriado à digitação. Após digitação, os arquivos passaram por um tratamento, eliminando os erros mais comuns encontrados, e os dados cadastrados com base na estrutura setorial de 1996 foram migrados e adequados à malha setorial 2000.

Este tratamento foi feito para os 1 058 municípios, com população em 1996 acima de 25 000 habitantes, o que resultou em um banco de dados preliminar para cada estado, contendo o resultado destes ajustes.

Esta versão, contendo ainda uma série de erros de transcrição e digitação, como a duplicação de logradouros devido à adequação à malha 2000, setores sem nenhum logradouro cadastrado previamente, entre outros, foi revisada e atualizada em campo pelo Supervisor na operação de Revisão da Base 2000 descrita no manual - Instruções para Revisão do Trabalho da Base Operacional Geográfica – TR 23.

A seguir, apresenta-se o relatório que foi utilizado para atualização em campo pelo supervisor.

Figura 7.8 - Formulário de Verificação do Cadastro de Segmentos de Logradouros



CADASTRO DE SEGMENTOS DE LOGRADOUROS - FORMULÁRIO DE VERIFICAÇÃO

UF: 14 RORAIMA
MUNICÍPIO: 00100 Boa Vista
DISTRITO: 05 BOA VISTA
SUBDISTRITO: 00
SETOR: 0001
SET./TIPO: 10

OP: 1 - INCLUSÃO 4 - ALTERAÇÃO E - EXCLUSÃO

14 . 00100 . 05 . 00 . 0001

O resultado deste trabalho é um banco de dados para aproximadamente 128 mil setores censitários, contendo cerca de 1 400 000 segmentos cadastrados.

Resultados finais da Base Operacional Geográfica 2000

As atividades de criação da Base Operacional Geográfica do Censo 2000 desenvolveram-se entre maio de 1997 e julho de 2000. A seguir, são apresentados os totais de setores de 2000 segundo o primeiro critério do código de formação (manutenção, agregação ou subdivisão e área alagada); e segundo a situação (urbana ou rural) e o tipo (não-especial ou especial).

Tabela 7.3 - Setores de 2000, segundo primeiro critério de formação

Primeira posição		Total de setores
Código	Composição da área	
1	Área do setor 2000 constituída pela área do setor de 96	90 074
2	Área do setor 2000 constituída por parte do setor de 96	86 236
3	Área do setor 2000 constituída pela junção de setores de 96 inteiros	3 644
4	Área do setor 2000 constituída pela junção de partes de setores de 96	25 113
5	Área do setor 2000 constituída pela junção de área (s) de setor (es) de 96 inteiro (s) com partes de setor (es) de 96	10 733
6	Área do setor 2000 parcialmente alagada	7
0	Área do setor totalmente alagada	4

Fonte: IBGE, Diretoria de Geociências, Departamento de Estruturas Territoriais

Tabela 7.4 - Setores de 2000, por situação, segundo a especificação do tipo do setor - 2000

Especificação do tipo do setor	Setores de 2000, por situação							
	Urbana			Rural				
	Cidade ou vila		Área urbana isolada	Aglomerado rural				Zona rural
	Área urbanizada	Área não urbanizada		De extensão urbana	Isolado			
					Povoado	Núcleo	Outros aglomerados	Exclusive aglomerado rural
Não especial	144 413	2 340	2 053	1 308	6 553	298	263	49 342
Especial de aglomerado subnormal (favelas e similares)	7 741	-	25	105	-	-	-	-
Especial de quartéis, bases militares, etc.	77	-	-	2	2	1	-	14
Especial de alojamentos, acampamentos, etc.	55	2	1	2	-	5	-	22
Especial de embarcações, barcos, navios, etc.	3	-	-	-	-	-	-	-
Especial de aldeia indígena	6	3	-	3	58	3	78	236
Especial de penitenciárias, colônias penais, presídios, etc.	158	10	3	-	2	-	-	46
Especial de asilos, orfanatos, conventos, hospitais, etc.	518	3	4	3	2	3	1	44

Fonte: IBGE, Diretoria de Geociências, Departamento de Estruturas Territoriais

7.3 Modernização da componente gráfica

Um dos mais importantes aperfeiçoamentos introduzidos na elaboração da Base Operacional Geográfica do Censo 2000 está associado à informatização da componente gráfica que substituiu, em grande medida, os lentos processos manuais até então utilizados.

7.3.1 Vertente urbana

A construção da base urbana digital do Censo 2000 teve por objetivo criar um acervo homogêneo de mapas urbanos, até então inexistente no IBGE, de forma padronizada e que viabilizasse atualização e reprodução, ágil e uniforme, das localidades brasileiras e dos setores censitários urbanos. Por não ser produtor de mapeamento em escala cadastral, o IBGE sempre se deparou com grandes dificuldades no preparo da base territorial urbana para uso em suas pesquisas.

Anteriormente, o que se fazia era a obtenção de cópias em papel destas bases a cada pesquisa. Isto, além de não construir um acervo utilizável, acarretava um enorme esforço na transcrição de todos os limites de estruturas territoriais e das atualizações cartográficas da pesquisa anterior para o planejamento da malha setorial da nova pesquisa. Como resultado, obtinha-se um gigantesco acervo em papel, impossível de ser armazenado corretamente (formatos diversos), de ser conservado ao longo do tempo (deteriorização da legibilidade) e, principalmente, de ser disseminado de forma padronizada e ágil.

O programa de trabalho para informatização da componente gráfica urbana do Censo 2000 foi estabelecido a partir da avaliação conjunta de questões como:

- as expectativas do IBGE e da comunidade usuária quanto à modernização e melhoria dos Mapas de Localidades Estatísticos (MLEs) e dos Mapas de Setores Urbanos (MSUs);
- a situação e disponibilidade de bases cadastrais nos órgãos produtores de mapeamento em escalas cadastrais, uma vez que não compete ao IBGE tal produção;
- a disponibilidade e o estágio de capacitação do pessoal na sede e nas unidades regionais do IBGE;
- a disponibilidade de recursos financeiros; e
- a obrigatoriedade de cumprimento de um cronograma retroativo, a partir da data de início da coleta do Censo 2000.

Esta avaliação dimensionou o salto tecnológico viável para a vertente urbana como sendo a construção de uma biblioteca CAD (*Computer Added Draft*)[1] padronizada, contendo Mapas de Localidades Estatísticos, atualizados para as cidades com população acima de 25 000 ou 50 000 habitantes, dada pela Contagem da População 1996. Este acervo gráfico, construído em nível de desenho digital, não deve ser visto como pronto para utilização em Sistemas de Informações Geográficas (SIG, GIS), uma vez que seu objetivo primeiro foi

[1] Desenho assistido por computador.

eliminar os inúmeros problemas com o preparo e a qualidade dos mapas dos setores urbanos para a coleta, bem como com o seu fornecimento posterior.

Ao programa de trabalho destinado a suprir as necessidades da operação de coleta através da modernização dos processos de desenho, atualização e impressão de MLEs e MSUs, denominou-se Programa Básico.

Entretanto, ainda que a biblioteca gráfica represente grande avanço para a operação censitária e inestimável contribuição ao acervo nacional de bases cadastrais digitais, deixa a desejar no que concerne a sua utilização para as análises e representações temáticas das informações territoriais e socioeconômicas resultantes. Isto se deve ao fato de que o mapeamento utilizado como base, obtido de diferentes órgãos, não é homogêneo quanto às suas características de geometria, escala de construção e unidade métrica, precisão de localização, referencial geodésico e sistema de projeção, apresentação (simbologia, fontes) etc., tendo ainda sido utilizados croquis com posicionamento aproximado, mapas oriundos de levantamentos topográficos e bases de restituição fotogramétrica, de acordo com a disponibilidade.

A construção da base urbana digital incluiu grandes etapas de trabalho como:

- levantamento do acervo urbano disponível (meio digital ou convencional);

- estabelecimento de acordos e parcerias com órgãos produtores e usuários que se dispuseram a fornecer este mapeamento em meio convencional e/ou digital;

- obtenção da documentação - mapas em papel e arquivos digitais;

- conversão para um padrão único, denominado IBGE – Censo 2000;

- impressão (plotagem) das folhas cadastrais das localidades (escala média 1:5 000);

- atualização dos limites das estruturas territoriais (perímetros urbanos legais, regiões administrativas e similares, bairros e similares, e outras) e planejamento da malha setorial de 2000;

- atualização cartográfica, principalmente de logradouros não constantes da base primária, e revisão da toponímia;

- desenvolvimento de sistemas customizados para implementação da produção digital, agilização do processo de capacitação de pessoal, e impressão (plotagem) semi-automática dos mapas das localidades e dos setores censitários urbanos;

- capacitação de cerca de 1 000 pessoas dos quadros permanente e temporário do IBGE;

- digitalização dos mapas (no caso do acervo adquirido em papel), das atualizações cartográficas e das malhas territoriais;

- geração de arquivos AGF (formato utilizado pelo *software* Atlas Gis) e Shape (formato utilizado pelo *software* Arc View), em sistemas de coordenadas geográficas e UTM, para aproximadamente 128 mil setores censitários 2000 (polígonos), das cidades dos municípios com população na Contagem da Populaçao 1996 superior a 25 mil habitantes, para subsidiar o processo de criação de áreas de expansão da amostra do Censo 2000 e para geração de produtos temáticos; e

- geração de arquivos em PDF (formato utilizado pelo *software* Adobe Acrobat) desses 128 mil setores visando facilitar a disseminação.

Toda a metodologia e processos operacionais adotados na construção deste acervo foram estabelecidos para a sustentação da operação censitária e das pesquisas internas do IBGE, nas suas atividades de planejamento, coleta, apuração e disseminação de resultados.

Foram desenvolvidos aplicativos na plataforma do *software* MicroStation, visando tanto à otimização da produção destes mapas a partir do mapa convencional (papel), quanto à adequação (conversão) dos mapas em meio digital, adquiridos através de convênios/parcerias.

Parcerias

O IBGE, além do seu próprio acervo cartográfico, contou com uma vasta rede de parceiros que produzem documentação cartográfica. A grande maioria é composta por órgãos de planejamento e institutos de pesquisa estaduais e municipais, incluindo-se também as concessionárias de serviços públicos, alguns institutos de pesquisa federais e empresas que atuam em geoprocessamento ou interessadas em desenvolver seus próprios sistemas de informações geográficas. Em anexo, apresenta-se a relação de produtores de documentação cartográfica, parceiros do IBGE.

7.3.2 Vertente rural

Reconhecendo-se a inviabilidade de se realizar a conversão total do mapeamento existente para o ambiente digital em formato CAD no período necessário às operações do Censo, o IBGE adotou como estratégia de operação a conversão para o formato raster[2] com atualização em formato vetorial[3], de cartas topográficas dos municípios com população acima de 25 000 habitantes, que perfaziam o montante de 1 058 municípios, abrangendo mais de 70% da população brasileira. Para o restante dos municípios seriam mantidas as bases elaboradas por processo convencional, com perspectiva de transformação para o ambiente digital de forma gradativa.

Entretanto, o formato digital híbrido (raster + vetor) acabou por ser adotado para todos os municípios que dispunham de base cartográfica. Com isto, foram produzidos aproximadamente 4 049, ou seja, 82% dos Mapas Municipais Digitais – MMDs e Mapas Municipais Digitais Estatísticos – MMDEs do total de 5 507 municípios brasileiros, além dos mapas de setores digitais rurais correspondentes. Por falta de mapeamento básico em escala adequada, 18%

[2] Formato raster - Estrutura matricial onde os dados são organizados em linhas e colunas, sendo associado a cada unidade de informação (pixel) uma cor.

[3] Formato vetorial - Os dados são organizados pelos seus descritores: linha, ponto, área.

dos municípios do Brasil não tiveram mapas municipais no formato digital para a coleta do Censo, e para estes municípios foram utilizados mapas municipais na forma convencional.

Diversas foram as ações implementadas para atingir as metas estabelecidas para a atividade de coleta no que concerne à cobertura cartográfica dos municípios. Dentre elas merecem destaque:

- a elaboração de mapas municipais no formato digital exigiu grande esforço para a formação de uma Base de Dados Raster. Foram convertidos para o meio digital todos os originais de cartas do Mapeamento Sistemático Brasileiro, nas escalas de 1:25 000 a 1:250 000, não só aquelas elaboradas pelo IBGE, como também por outros órgãos como a Diretoria do Serviço Geográfico do Exército – DSG. Desta conversão, resultou um grande volume de arquivos digitais no formato Raster, que após passarem por um tratamento digital da imagem, foram georreferenciados, tendo sido armazenados em um servidor do IBGE, no Rio de Janeiro, para utilização na elaboração dos mapas municipais;

- a atualização cartográfica de campo e em gabinete, utilizando como base o mapeamento sistemático topográfico, procurou garantir, pelo menos, a representação dos aglomerados populacionais, principais vias de acesso e de alguns estabelecimentos agropecuários;

- a realização do mapeamento das áreas sem cobertura topográfica, mesmo que com alternativas preliminares ou provisórias;

- a geração e atualização de cadastros correlatos (produtores e usuários, topônimos, insumos, entre outros); e

- a edição das malhas municipais e intramunicipais dos 5 507 municípios, além da malha setorial com incorporação de perímetros urbanos.

As demandas dos usuários por produtos digitais e a necessidade de melhoria geral da base cartográfica deflagraram um processo de transformação na produção cartográfica municipal do IBGE. Esta demanda exigiu a contratação de serviço de consultoria para o desenvolvimento de um sistema que atenderia, a curtíssimo prazo, às solicitações e às necessidades. O Sistema de Cartografia Semi-Automatizada – SISCART foi desenvolvido para elaborar mapas municipais, Mapas Municipais Estatísticos – MMEs e Mapas de Setores Rurais – MSRs, de forma semi-automática, com a mínima intervenção do operador.

O sistema passou por diversas fases de desenvolvimento, até se adaptar às especificidades do projeto Censo 2000. A elaboração de mapas municipais digitais, como finalidade principal do SISCART, foi a fase que consumiu a maior parte do tempo previsto para o seu desenvolvimento, baseado na forma convencional de elaboração dos mapas municipais, ou seja, a partir da compilação das folhas topográficas que formam a base topográfica para o lançamento do polígono municipal resultando no mapa municipal.

Os produtos do SISCART foram gerados por processos semi-automáticos de associação de arquivos digitais nos formatos Raster e Vetor e bancos de dados. Para que fosse possível a automatização dos processos, foi necessário que todo o mapeamento existente fosse convertido para o formato digital, utilizando-se, no

caso o Raster por ser mais rápido. O acervo do IBGE e o de autoria de outros órgãos foram convertidos, formando uma mapoteca com arquivos Raster que passaram para a próxima fase, o georreferenciamento.

Como já foi dito, o IBGE, além do seu próprio acervo cartográfico, contou com uma vasta rede de parceiros, que disponibilizaram os originais das folhas topográficas de sua autoria. A grande dificuldade enfrentada foi a localização de alguns dos originais dos documentos cartográficos que seriam convertidos para o formato digital via *scanner*. O extravio ou a perda de alguns originais fez com que a conversão se desse a partir da folha topográfica impressa, com a utilização de *software* específico e de técnicos capacitados para a obtenção dos arquivos digitais.

A conversão de todos os originais de cartas topográficas existentes nas escalas entre 1:25 000 e 1:250 000 formou o acervo digital do Departamento de Cartografia, composto pelos arquivos Raster das folhas topográficas, divididos em categorias de informação cartográfica, correspondendo aos elementos físicos naturais e artificiais do terreno: rios e massas d'água, sistema viário, localidades, obras e edificações e também composto por arquivos no formato Vetorial, representando as linhas de limites municipais, distritais e de setores.

Basicamente, a metodologia de elaboração dos mapas municipais digitais consistiu na junção dos arquivos Raster, das folhas topográficas para a composição da área municipal, sobre a qual foram lançadas as linhas de limites. Após a incorporação das atualizações e verificação da consistência das linhas de limites foi feito o recorte do município e, sobre ele, foram lançadas as molduras para o acabamento final, que fornece as informações sobre as representações cartográficas utilizadas.

As informações constantes dos mapas municipais visam à adequada representação do espaço geográfico no limite municipal e seu entorno, contendo os elementos físicos constantes nos arquivos Raster utilizados na sua elaboração, bem como coordenadas geográficas e UTM e linhas de limites intramunicipais.

O SISCART utilizou aplicativos gráficos computacionais, operando sobre a plataforma Windows, tendo sido desenvolvido em Visual Basic gerenciando *softwares*, tais como: MICROSTATION, IRASB, MGE, IPLOT associados a bancos de dados ACCESS, com estrutura de relacionamentos em ambiente CLIENTE/SERVIDOR.

As fases da elaboração dos MMDs, MMDEs e MSRs são as seguintes:

1. elaboração da base cartográfica municipal;

2. incorporação de informações atualizadas em campo com receptor GPS;

3. digitalização das malhas distritais e setoriais;

4. produção de MMDs, MMDEs e MSRs; e

5. plotagem dos MMDs, MMDEs e MSRs.

O padrão SISCART de produção e armazenamento das informações digitais foi adotado para melhor atender às demandas e facilitar a inclusão de novas informações geradas interna e externamente ao IBGE. Com a aplicação de geotecnologias, foi possível a disponibilização de bases cartográficas digitais com resultados estatísticos georreferenciados.

A tabela, a seguir, apresenta o total de municípios, cujas bases foram prioritariamente definidas para conversão em formato digital (coluna 3) e o quantitativo de municípios que tiveram uma melhoria geral da base cartográfica municipal, inclusive com a elaboração de mapas municipais digitais (colunas 4 e 5).

Tabela 7.5 - Quantitativo de mapas municipais digitais elaborados, por classes de população do município, segundo as Grandes Regiões - 2000

Grandes Regiões	Total de Municípios	Municípios com população acima de 25 000 habitantes (coluna 3)				Municípios com população entre 10 000 e 25 000 habitantes (coluna 4)				Municípios com população abaixo de 10 000 habitantes (coluna 5)			
		Total	MMD	MMDE	MSR	Total	MMD	MMDE	MSR	Total	MMD	MMDE	MSR
Brasil	**5 507**	**1 058**	**1 058**	**1 058**	**26 648**	**1 368**	**1 368**	**1 368**	**18 878**	**2 640**	**2 640**	**2 640**	**14 372**
Norte	313	79	79	79	2 033	81	81	81	936	172	172	172	778
Nordeste	1 780	392	392	392	9 542	727	727	727	9 360	662	662	662	3 813
Sudeste	1 630	364	364	364	8 261	148	148	148	4 220	832	832	832	3 823
Sul	1 089	154	154	154	5 610	257	257	257	2 983	706	706	706	4 534
Centro-Oeste	585	69	69	69	1 202	155	155	155	1 379	268	268	268	1 424

Fonte: IBGE, Diretoria de Geociências, Departamento de Cartografia.

Referências

XI RECENSEAMENTO geral do Brasil: manual de delimitação dos setores de 2000. Rio de Janeiro: IBGE, Diretoria de Geociências, 1998.

BASE territorial do censo 2000: manual de utilização do sistema BOG 2000. Rio de Janeiro: IBGE, Diretoria de Geociências, 1998.

BASE territorial do censo 2000: vertente urbana: informatização da componente gráfica: referências para a compatibilização dos arquivos da base territorial e para a construção e impressão dos mapas de setores (MSU). Rio de Janeiro: IBGE, Diretoria de Geociências, 1999.

BASE territorial do censo 2000: vertente urbana: informatização da componente gráfica: referências para conversão e construção de mapas de localidades no padrão IBGE_Censo 2000 (versão 5.0). Rio de Janeiro: IBGE, Diretoria de Geociências, 1998.

BASE territorial do censo 2000: vertente urbana: referências para impressão dos mapas de localidades e dos mapas de setores urbanos 2000. Rio de Janeiro: IBGE, Diretoria de Geociências, 2001.

BASE territorial do censo 2000: vertente urbana: referências para instalação e utilização do sistema de impressão e atualização do cadastro de segmentos de logradouros. Rio de Janeiro: IBGE, Diretoria de Geociências, 2001.

CENSO demográfico 2000: instruções para revisão do trabalho da base operacional geográfica - TR 23. Rio de Janeiro: IBGE, 2000.

MANUAL da malha censitária. Rio de Janeiro: IBGE, Diretoria de Geociências, 2001.

MANUAL de atualização cartográfica. Rio de Janeiro: IBGE, Diretoria de Geociências, 2000.

MANUAL de atualização de bases cartográficas para a coleta do censo 2000. Rio de Janeiro: IBGE, Diretoria de Geociências, 1998.

MANUAL de controle de qualidade. Rio de Janeiro: IBGE, Diretoria de Geociências, 2000.

MANUAL de edição da malha municipal. Rio de Janeiro: IBGE, Diretoria de Geociências, 2000.

MANUAL de elaboração do mapa. Rio de Janeiro: IBGE, Diretoria de Geociências, 2000.

MANUAL de georeferenciamento. Rio de Janeiro: IBGE, Diretoria de Geociências, 2000.

MANUAL de instalação do SISCART. Rio de Janeiro: IBGE, Diretoria de Geociências, 2000.

MANUAL de reprodução. Rio de Janeiro: IBGE, Diretoria de Geociências, 2001.

MANUAL para utilização de cartas expeditas. Rio de Janeiro: IBGE, Diretoria de Geociências, 2001.

MANUAL de georeferenciamento. Rio de Janeiro: IBGE, Diretoria de Geociências, 2000.

MANUAL de instalação do SISCART. Rio de Janeiro: IBGE, Diretoria de Geociências, 2000.

MANUAL de reprodução. Rio de Janeiro: IBGE, Diretoria de Geociências, 2001.

MANUAL para utilização de cartas expeditas. Rio de Janeiro: IBGE, Diretoria de Geociências, 2001.

Anexo

Relação de Parceiros

Prefeitura Municipal em todos os Municípios
Pará CODEM - Companhia de Desenvolvimento e Administração de Belém COHAB - Companhia de Habitação TELPA - Telecomunicações do Estado do Pará
Amapá Secretaria do Meio Ambiente
Maranhão CAEMA - Companhia de Água e Esgoto do Maranhão
Piauí Água e Esgoto do Piauí S/A TELEMAR (ex. TELEPISA - Telecomunicações do Estado do Piauí)
Ceará SEINFRA (ex SEDURB - Secretaria de Desenvolvimento Urbano) IPLAN - Instituto de Planejamento
Rio Grande do Norte CAERN - Cooperativa de Água e Esgoto COSERN - Companhia de Serviços Elétricos do Rio Grande do Norte TELEMAR (ex. TELERN - Telecomunicações do Estado do Rio Grande do Norte)
Paraíba CAGEPA - Cia. De Água e Esgoto da Paraíba SAELPA - S/A de Eletrificação da Paraíba Secretaria de Planejamento do Município de João Pessoa TELEMAR (ex. TELPA - Telecomunicações da Paraíba S/A)
Pernambuco FIDEM - Fundação do Desenvolvimento da Região Metropolitana de Recife
Sergipe Secretaria Estadual de Educação
Bahia Secretaria Estadual de Educação CONDER - Companhia de Desenvolvimento da Região Metropolitana de Salvado
Minas Gerais CEMIG - Centrais Elétricas de Minas Gerais, Empresa Cataguases Leopoldinense - Companhia de Força e Luz PRODABEL TELEMAR
Espírito Santo TELEMAR (ex. TELEST - Telecomunicações do Estado do Espírito Santo)
Rio de Janeiro Águas de Juturnaíba CEFET CERJ Fundação CIDE - Centro de Informações e Dados do Rio de Janeiro IBAM IPP - Instituto Pereira Passos LIGHT TELEMAR

Relação de Parceiros

São Paulo Empresa Geograph Informática e Serviços Ltda
Paraná Empresa Serviço Social Autônomo PARANACIDADE IPARDES - Instituto Paranaense de Desenvolvimento Econômico e Social IPPUC - Instituto de Planejamento e Pesquisa Urbana de Curitiba
Santa Catarina CASAN - Companhia de Água e Saneamento ENGEVIX - Engenharia S/A IPUF- Instituto de Planejamento Urbano de Florianópolis IPUJ - Instituto de Planejamento Urbano de Joinville MPB Engenharia Ltda. SAMAE - Serviço Autônomo Municipal de Água e Esgoto SERENCO - Serviço de Engenharia Consultiva Ltda. (Curitiba – PR)
Mato Grosso do Sul IDATERRA - Instituto de Desenvolvimento Agrário, Assistência Técnica e Extensão do Mato Grosso do Sul INCRA - Instituto Nacional de Reforma Agrária SANESUL - Empresa de Saneamento do Mato Grosso do Sul TOPOSAT Ltda - Empresa de Topografia
Mato Grosso TELEMAT - Companhia de Telefonia do Estado de Mato Grosso
Goiás CELG - Centrais Elétricas de Goiás FNS - Fundação Nacional de Saúde
Distrito Federal CODEPLAN - Companhia de Desenvolvimento e Planejamento do Distrito Federa SEDUH - Secretaria de Estado de Desenvolvimento Urbano e Habitação (ex IDAHAB) SUDUR - Sub-Secretaria de Urbanismo e Preservação (ex IPDF)

8 Treinamento

O programa de treinamento do Censo Demográfico 2000 teve como objetivo assegurar a uniformidade no desempenho de todos os participantes das atividades censitárias, provendo os treinandos de instrumentos que os habilitassem a preencher e/ou orientar o preenchimento correto dos formulários de coleta e a atuar como instrutores e multiplicadores de conhecimentos sobre o Censo.

Para tanto, elaborou-se um projeto que estruturasse essa importante etapa da operação censitária, de forma a integrar não só os conteúdos técnicos – manuais de coleta – mas todas as áreas de trabalho. O projeto objetivava, ainda, a incorporação das instruções administrativas, dos sistemas informatizados e da Base Operacional e Geográfica. Buscava também, aprimorar os instrumentos de treinamento e as técnicas a serem aplicadas, mediante o uso de ferramentas modernas de ensino a distância, a fim de assegurar a transmissão homogênea das instruções, evitando as distorções comuns nos processos de treinamento reproduzidos em cadeias sucessivas e crescentes.

Para elaborar o projeto, foi constituído o Grupo de Trabalho do Projeto Treinamento do Censo 2000 – GTT. Sob a coordenação de consultoria externa, o grupo, composto por oito técnicos das diversas áreas mais diretamente envolvidas no assunto, Diretoria de Pesquisas – DPE, Diretoria de Geociências – DGC, Centro de Documentação e Disseminação de Informações – CDDI, Escola Nacional de Ciências Estatísticas – ENCE, Coordenação de Acompanhamento e Controle Operacional dos Censos – COC e Coordenação de Treinamento – COT – elaborou o documento *Projeto de treinamento do censo 2000* (1999) e acompanhou a

sua implementação, fazendo, ao final, a avaliação, que resultou no *Relatório de avaliação do programa de treinamento do censo demográfico 2000* (2000). A coordenação central do treinamento foi compartilhada com a COC e a DPE.

O projeto, o relatório de avaliação, o material didático – exceto as fitas de vídeo – e os demais instrumentos utilizados pelo programa de treinamento constam no CD-ROM, anexo a esta publicação.

8.1 Estratégia

O Projeto de Treinamento do Censo 2000 estabeleceu o desenvolvimento das seguintes etapas: o Curso de Capacitação Didático-pedagógica e o Treinamento técnico-operacional, este último com a seguinte estruturação:

- técnico – abrangendo os manuais técnicos, ligados diretamente à coleta de dados; e
- operacional ou administrativo/gerencial – abordando os manuais referentes às atividades de apoio à coleta – operação dos sistemas informatizados, procedimentos administrativos e revisão da Base Operacional e Geográfica.

O Curso de Capacitação Didático-pedagógica, cuja coordenação e execução ficou a cargo da ENCE, teve como principal objetivo desenvolver, nos futuros instrutores censitários, habilidades e atitudes necessárias à sua atuação em sala de aula. Previa, ainda, a capacitação desses instrutores na aplicação das técnicas selecionadas e na utilização dos recursos instrucionais do treinamento do Censo.

O treinamento técnico – para a coleta de dados - foi desenvolvido através de um processo de capacitação em cadeia, mediante a formação de grupos/equipes constituídos por servidores do IBGE e pelo pessoal contratado para as funções censitárias de Agentes Censitários Municipais, de Agentes Censitários Supervisores e de Recenseadores.

O processo de treinamento teve início, no Rio de Janeiro, com a formação de um primeiro grupo de instrutores da DPE e prosseguiu, em cadeia, através dos seguintes grupos e etapas de execução:

- Grupo 1 – composto por servidores do quadro permanente, da sede do IBGE e das Unidades Regionais; etapa centralizada em Guarapari/ES, para formar os instrutores do Grupo 2;
- Grupo 2 – composto por servidores das URs – Coordenadores de Área, de Subárea e outros servidores; etapa descentralizada em 14 pólos regionais, para formar os instrutores do Grupo 3;
- Grupo 3 – composto por chefes de agências e outros técnicos que não haviam participado dos grupos anteriores; etapa descentralizada em cada UR, para formar os instrutores dos grupos subseqüentes;
- Grupo 4 – composto por ACMs e ACSs; etapa descentralizada, realizada em pólos, em geral constituídos por municípios de uma mesma agência, em cada UR ; e
- Recenseadores – composto pelos Recenseadores; etapa realizada em todos os municípios do País.

O treinamento operacional, para a capacitação administrativa e gerencial, teve como objetivo preparar os treinandos para o desempenho de funções gerenciais durante a coleta e para a realização dos trabalhos preparatórios da operação censitária, relativos à organização dos postos de coleta, à revisão da BOG e ao uso dos sistemas administrativo e gerencial.

O treinamento foi iniciado com a preparação das equipes de instrutores da DE, DI e DGC, áreas responsáveis pelos conteúdos que seriam ministrados mais tarde. Após esta etapa inicial, a instrução prosseguiu com as demais clientelas do programa até atingir o último nível gerencial, composto pelos ACMs e ACSs,

Para os Grupos 1, 2 e 3 a instrução ocorreu imediatamente após o treinamento técnico para a coleta de dados, sendo realizada em regime de internato para os dois primeiros grupos, e em serviço para o último. Os ACMs e ACSs, público-alvo do G4, receberam as instruções sobre os aspectos operacionais logo após a contratação, sendo os ACMs em regime de internato, e os ACSs em serviço.

A seguir, são apresentadas as equipes por categoria e o número de pessoas envolvidas no Treinamento técnico-operacional.

Tabela 8.1 - Número de instrutores e treinandos, segundo os grupos do treinamento - 2000

Grupos do treinamento	Instrutores	Treinandos
DPE (técnico)	2	16
DE/DI/DGC (operacional)	5	27
G1	50	220
G2	98	1 197
G3	82	838
G4	2 363	30 683
Recenseadores	17 667	189 270

Fonte: Relatório de avaliação do programa de treinamento do censo demográfico 2000. Rio de Janeiro: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, 2000. 44 p.

Esta estratégia de capacitação em cadeia vem sendo adotada pelo IBGE para os treinamentos censitários por permitir atingir, em curto período de tempo, um quantitativo elevado de pessoas, em vários níveis e, em diferentes localizações geográficas. Fundamental, no entanto, para o Censo 2000, foi minimizar as distorções e perdas de conteúdo ao longo dos sucessivos repasses de conhecimento, que ocorreram em maior ou menor grau nos censos anteriores.

Nesse sentido, com o propósito de neutralizar as disfunções decorrentes da estratégia adotada, foram definidas e implementadas as seguintes medidas:

- criação de um grupo especial de instrutores, todos servidores do quadro permanente de pessoal do IBGE, das unidades, da sede e das URs, treinados nos Grupos 1 e 2 e, especialmente, capacitados para atuar de forma direta e permanente, ainda que não exclusiva, nas várias etapas do treinamento descentralizado;

- realização do Curso de Capacitação Didático-pedagógica, em caráter obrigatório, para todos os instrutores do grupo especial de instrutores, com a finalidade de propiciar o embasamento necessário a sua atuação em sala de aula;

- realização do treinamento para os servidores do Grupo 2, em pólos regionais, de modo a possibilitar o treinamento direto, pelos instrutores formados no Grupo 1, a um número maior de servidores das Urs;

- utilização do recurso do vídeo-aula como forte apoio às atividades do treinamento, para permitir a visualização e a homogeneização dos assuntos tratados e para assegurar o desenvolvimento uniforme das atividades de ensino/aprendizagem em cada nível a ser treinado;

- elaboração dos manuais técnicos com orientação didático-pedagógica, utilizando linguagem coloquial, de fácil leitura e compreensão, e com estruturação e editoração atraentes e motivadoras;

- criação da função de Coordenador Estadual de Treinamento nas URs, na estrutura de cargos comissionados do Censo Demográfico 2000, com a atribuição de coordenar a operacionalização dos treinamentos realizados nas respectivas unidades;

- elaboração de manuais sobre os aspectos administrativos e de informática, visando abranger, com a maior amplitude possível, todos os aspectos normativos, operacionais e gerenciais da operação censitária;

- incorporação das atividades administrativas e gerenciais relativas à coordenação e supervisão de coleta no processo de treinamento técnico;

- realização de um treinamento destinado aos ACMs e ACSs, logo após a sua contratação, dois meses antes da coleta, sobre os manuais gerenciais da coleta e sobre o manual de Revisão da Base Operacional e Geográfica ; e

- inclusão, nos processos seletivos de ACM, ACS e Recenseador, da prova específica sobre os conceitos básicos do Censo Demográfico 2000, objeto da apostila distribuída no ato da inscrição do candidato.

8.2 Metodologia e execução

Nesta seção são apresentados os conteúdos programáticos, as cargas horárias e as técnicas de ensino aplicadas em cada um dos blocos que compuseram o Projeto Treinamento do Censo 2000 – o Curso de Capacitação Didático-pedagógica, e as duas partes distintas do treinamento técnico-operacional.

Curso de Capacitação Didático-pedagógica

É bom reafirmar que esse curso consistiu em etapa subsidiária do processo de seleção dos candidatos ao grupo especial de instrutores, que receberam, posteriormente, o treinamento técnico-operacional.

O conteúdo programático do curso, com carga horária total de 40 horas, foi desenvolvido no período de cinco dias e distribuído nos seguintes tópicos:

- A importância do treinamento do Censo 2000, identificando o papel do instrutor;

- O processo ensino-aprendizagem;

- Objetivos de ensino X alcance dos resultados;

- Relações interpessoais instrutor-treinandos;

- O processo de comunicação, aprendendo a dar e receber *feedback*;

- Administrando o medo de falar em público, valorização da postura corporal, aprendendo a estimular grupos;

- Técnicas de ensino e recursos instrucionais;

- Como operar TV e vídeo, explorando o vídeo-aula do treinamento do Censo 2000, Preparação da prática de ensino; e

- Prática de ensino e avaliação do curso.

O curso foi realizado no período de abril de 1999 a março de 2000 para grupos de, em média, 15 treinandos e desenvolvido, em um primeiro momento, para os servidores lotados na sede do IBGE e, logo após, para os servidores das Unidades Regionais, nas respectivas Unidades da Federação – UFs. Foram treinados 1 436 servidores do IBGE, em um total de 99 turmas, utilizando 21 instrutores.

Treinamento técnico

Buscando assegurar a fixação dos conteúdos transmitidos e proporcionando aos treinandos segurança na operacionalização dos conceitos e procedimentos da coleta, o treinamento técnico do Censo 2000 foi desenvolvido através de atividades de auto-instrução e da realização de cursos presenciais, com o apoio de um dispositivo de capacitação, que contemplou a utilização de vídeos, manuais e outros materiais impressos. Na orientação didática foram aplicadas técnicas de dinâmica de grupo, simulação de casos e exercícios, visando estimular a participação ativa dos treinandos em seu processo de aprendizagem, de modo a propiciar segurança na aplicação dos procedimentos de coleta e na transmissão dos conhecimentos.

A metodologia adotada foi desenvolvida através de roteiros de aulas, referentes às unidades de ensino dos conteúdos que seriam ministrados para cada clientela.

Com o fim de assegurar o desenvolvimento uniforme das atividades de ensino/aprendizagem, o roteiro de cada aula forneceu aos instrutores a seqüência do processo, conjugando conteúdo, estratégias de ensino e exercícios de verificação da aprendizagem, elementos-chave em qualquer prática pedagógica.

O Manual do Recenseador e os demais manuais técnicos, com os instrumentos de coleta, constituíram o conteúdo dos recursos utilizados.

Para todos os treinandos o aprendizado foi iniciado com a auto-instrução, cuja finalidade foi familiarizá-los, antes de entrar em sala de aula, com os conceitos, critérios e procedimentos que seriam adotados no trabalho de coleta de dados, de modo que pudessem tirar o maior proveito possível do treinamento presencial e, mais facilmente, aprender e fixar o conteúdo ministrado.

Nesta fase do treinamento, com duração de cinco dias, o treinando deveria ler o material recebido e fazer os exercícios propostos. Aos ACMs, ACSs e Recenseadores, que, em suas residências, dispusessem de televisão com antena parabólica, era recomendado assistir, também, à projeção da fita de vídeo - Censo 2000: vamos contar o Brasil –,transmitida no canal da TV Escola, do Ministério da Educação, nos horários determinados pela emissora.

A participação dos Recenseadores no curso presencial, conforme previsto no edital do processo seletivo, estava condicionada à apresentação, no primeiro dia de aula, das respostas ao Teste Inicial, que deveria ser feito após a leitura do Manual do Recenseador, durante a auto-instrução.

O conteúdo programático do treinamento para a coleta de dados, para toda a clientela do programa, abrangeu os seguintes tópicos:

- O recenseador e seus instrumentos de coleta – uma aula;

- O recenseador e a cobertura do seu percurso – duas aulas;

- O recenseador e o entrevistado: uma parceria necessária – uma aula;

- O recenseador e o preenchimento dos instrumentos – oito aulas; e

- Questionário Básico e da Amostra – quinze aulas.

Com exceção dos Recenseadores, para as demais clientelas, o conteúdo programático abrangeu, ainda, os temas a seguir relacionados:

- Funções e atribuições do supervisor: antes, durante e depois dos trabalhos da coleta – cinco aulas;

- O trabalho de coordenação das atividades da coleta – duas aulas; e

- Orientações didático-pedagógicas – três aulas.

De acordo com o conteúdo programático, a carga horária do treinamento dos Recenseadores foi de 27 horas, desenvolvida no período de quatro dias. Para os treinandos servidores do IBGE e para os ACMs e ACSs, a carga horária do curso foi maior, com vistas a contemplar as instruções específicas de suas funções de acompanhamento e controle da coleta, e de instrutor do Censo.

A cada turma formada para o treinamento correspondeu a atuação de dois instrutores que, previamente, deveriam estabelecer a distribuição das aulas entre si.

Ao término do treinamento, para todos os grupos de treinandos, foi aplicado um teste de avaliação de aprendizagem do conteúdo técnico, o Teste Final, com o objetivo de verificar se o treinando estava capacitado a desempenhar suas atribuições durante a coleta.

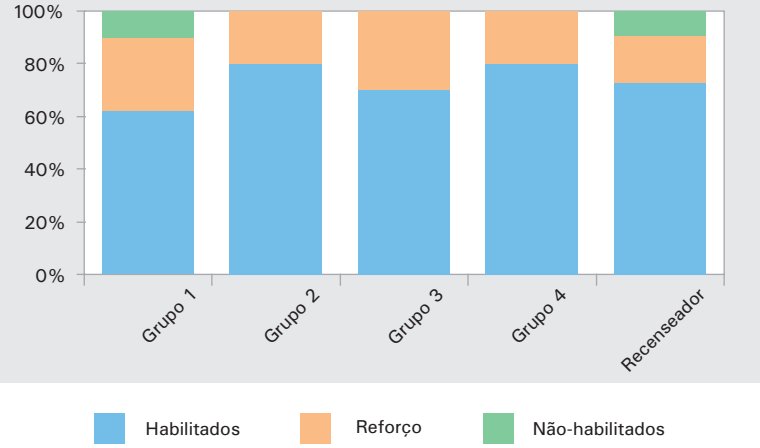
Alcançando um resultado acima de 60% de acertos, o treinando era considerado habilitado ao desempenho das atribuições; entre 40% e 60%, indicado para reforço; e abaixo de 40%, classificado como não-habilitado. O reforço de treinamento consistiu no esclarecimento dos erros cometidos ou no retreinamento.

No caso específico da escolha dos técnicos que atuaram como instrutores no Grupo 2, selecionados entre os treinandos habilitados do Grupo 1, levou-se em conta os graus obtidos no Curso de Capacitação Didático-pedagógica e no teste de avaliação da aprendizagem ou de conteúdo técnico, devidamente ponderados. Os pesos de ponderação foram 30 e 70, respectivamente.

No treinamento dos Recenseadores, o resultado do Teste Final indicava se o treinando seria ou não contratado para o Censo 2000.

O resultado da avaliação da aprendizagem consta do gráfico seguinte

Gráfico 8.1 - Avaliação da aprendizagem, segundo as etapas do treinamento



Fonte: Relatório de avaliação do programa de treinamento do censo demográfico 2000. Rio de Janeiro: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, 2000. 44 p.

Treinamento operacional ou administrativo/gerencial

O treinamento referente aos manuais gerenciais – Manual do Agente Censitário Municipal – CD 1.11, Manual do Usuário do Sistema de Indicadores Gerenciais da Coleta de Dados – TR 24 e Manual do Sistema de Administração de Pessoal – AD 01 – e ao manual de Instruções para Revisão do Trabalho da Base Operacional Geográfica – TR 23 foi ministrado para os treinandos dos Grupos 1 e 2 e para os ACMs obedecendo às mesmas orientações e procedimentos didáticos adotados no treinamento técnico para a coleta de dados. Entretanto, para os servidores do IBGE, treinandos do Grupo 3, e para os ACSs, optou-se pela execução do treinamento em serviço, devido ao custo elevado que implicaria o deslocamento deste pessoal para pólos de treinamento.

A carga horária deste treinamento foi de 16 horas, assim distribuídas:

- Manual do ACM – quatro horas;
- Manual do SIGC – três horas;
- Manual do SAP – quatro horas; e
- Manual de Revisão da BOG – cinco horas.

Para os ACMs, lotados em postos de coleta não-informatizados ou em municípios com menos de 25 000 habitantes, onde não seria realizado o trabalho de revisão do Cadastro de Segmentos de Logradouros, foi reduzida a carga horária do treinamento relativa aos assuntos dos manuais do SIGC e da BOG.

8.3 Recursos didáticos

Para o treinamento técnico-operacional destinado à coleta de dados do Censo, foram preparados os seguintes recursos:

- Manual do Instrutor – TR 01 – contendo os roteiros das aulas e todas as orientações necessárias ao trabalho do instrutor;

- Álbum Seriado – TR 14 – contendo um conjunto de cartazes referentes aos temas das aulas;

- Vídeo-aula – TR 25 – compreendendo uma fita de vídeo estruturada em 16 módulos, sendo um para cada tema do conteúdo técnico-operacional ministrado, e um sobre o conteúdo das aulas didático-pedagógicas;

- Roteiro de Estudo – TR 02 – contendo os exercícios de fixação voltados para auxiliar o treinando na leitura do Manual do Recenseador durante a auto-instrução;

- Teste Inicial – TR 03 – contendo as questões referentes aos conceitos e procedimentos contidos no Manual do Recenseador;

- Caderno de Exercícios – TR 04 – compreendendo um conjunto de exercícios destinados à fixação dos conceitos e procedimentos definidos para o Censo, bem como à forma de preenchimento dos instrumentos de coleta;

- Instrumentos Ampliados – cópias ampliadas das folhas de coleta – TR 15 e das folhas de domicílio coletivo – TR 16, utilizadas na correção, em grupo, dos exercícios sobre o preenchimento destes instrumentos;

- Teste Final – TR 06 e TR 07 – abrangendo os modelos A e B: sendo o modelo A para os treinandos até o nível de Supervisor e o modelo B para os Recenseadores, ambos os testes apresentando situações para o preenchimento dos instrumentos de coleta;

- Avaliação do Treinamento – formulários TR 11, TR 12 e TR 13 contendo as informações sobre as atividades do treinamento a serem avaliadas, respectivamente, pelos treinandos, instrutores e observadores, e

- Roteiros das Aulas – TR 17 – contendo as orientações necessárias à condução de treinamentos rápidos (de 12 horas) para Recenseadores recrutados, emergencialmente, após o início da coleta.

Para o treinamento referente aos manuais gerenciais e de revisão da BOG, foram, também, preparados:

- Manual do Instrutor – TR 18 – contendo os roteiros das aulas sobre os manuais gerenciais e de revisão da BOG;

- Caderno de Exercícios – TR 19 – contendo os exercícios sobre a revisão da BOG;

- Álbum Seriado – TR 20 – contendo um conjunto de cartazes que abordavam o conteúdo das aulas sobre os manuais gerenciais e de revisão da BOG; e

- Avaliação do Treinamento – formulários TR 21, TR 22, e TR 31 contendo as informações sobre as atividades do treinamento a serem avaliadas, respectivamente, pelos treinandos, instrutores e observadores.

Todos os manuais técnicos utilizados no programa de treinamento do Censo 2000 também receberam tratamento didático-pedagógico, a saber:

- Manual do Recenseador – CD 1.09;
- Manual do Supervisor – CD 1.10;
- Manual do Agente Censitário Municipal – CD 1.11;
- Manual de Coordenação Técnica – CD 1.12;
- Manual do Usuário do Sistema de Indicadores Gerenciais da Coleta de Dados – TR 24;
- Manual do Sistema de Administração de Pessoal – AD 01; e
- Instruções para Revisão do Trabalho da Base Operacional Geográfica – TR 23.

8.4 Acompanhamento e avaliação

A forma como foram planejados os trabalhos de acompanhamento e avaliação do treinamento possibilitou a participação, nessas atividades, de boa parte do corpo técnico envolvido na operação censitária.

Assim, tanto os servidores do quadro permanente como os contratados, ao desempenharem suas funções no processo de treinamento, como treinandos, instrutores ou observadores, contribuíram, dependendo de suas competências, para o acompanhamento e/ou avaliação do treinamento.

8.4.1 Avaliação dos observadores

Ao longo do processo de treinamento, a Coordenação Central do Treinamento manteve um canal aberto de comunicação com os coordenadores de treinamento dos estados, via correio eletrônico, de modo a atendê-los na solução dos problemas e dúvidas surgidas.

Assim, complementando as orientações contidas nas normas e rotinas, os boletins *Informativo treinamento censo 2000*, emitidos quando necessário, transmitiam aos Coordenadores Estaduais de Treinamento as orientações sobre os procedimentos para a operacionalização do processo descentralizado.

Por sua vez, os relatórios de acompanhamento e avaliação do treinamento, enviados pelos Coordenadores, informavam sobre cada etapa concluída, o que permitiu a adoção de medidas e providências para solucionar as disfunções indicadas. Isso aconteceu, por exemplo, quando foi ampliada a carga horária do treinamento, a partir do Grupo 3, a fim de redimensionar o tempo destinado à instrução do preenchimento da Folha de Coleta.

Ainda como parte do trabalho, nas etapas descentralizadas, até o treinamento de ACM, em todas as unidades regionais, havia técnicos da Coordenação Central de Treinamento acompanhando o processo. Os treinamentos de Supervisores e Recenseadores contaram sempre com o acompanhamento de técnicos das próprias unidades regionais e, sempre que possível, da Coordenação Central de Treinamento, na qualidade de observadores externos, conforme discriminado abaixo:

Tabela 8.2 - Número de observadores e treinandos observado - 2000

	Número
Observadores	214
Turmas observadas	893
Treinandos observados	22 325
Instrutores observados	1 789

Fonte: Relatório de avaliação do programa de treinamento do censo demográfico 2000. Rio de Janeiro: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, 2000. 44 p.

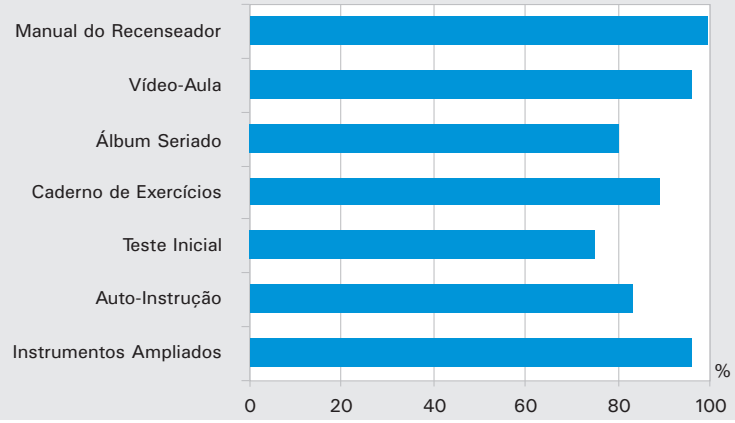
Pelo TR 13 Avaliação do Treinamento – Observador, foram avaliados os seguintes aspectos do treinamento: recursos didáticos utilizados; competência técnica do instrutor; desempenho dos treinandos; e sistemática do treinamento.

A seguir, são apresentados alguns dos resultados do trabalho de observação do treinamento técnico-operacional:

a) quanto aos recursos didáticos elaborados

O Gráfico 8.2 mostra a proporção de Observadores que consideraram os recursos didáticos cumprindo sua finalidade de facilitar o processo de ensino/aprendizagem.

Gráfico 8.2 – Resultado da avaliação dos recursos didáticos

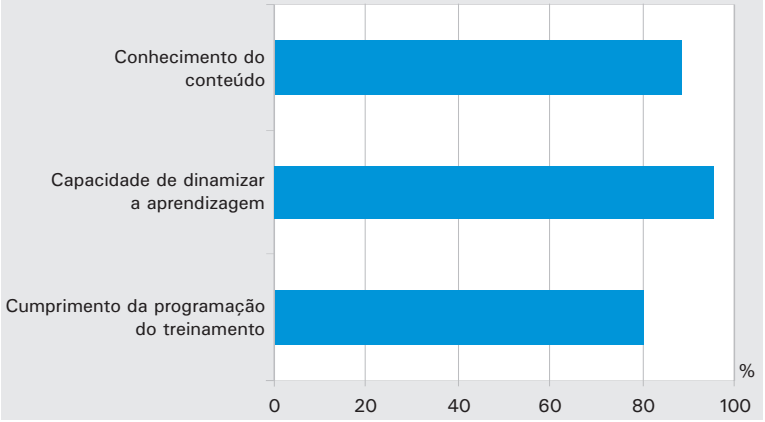


Fonte: Relatório de avaliação do programa de treinamento do censo demográfico 2000. Rio de Janeiro: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, 2000. 44 p.

b) quanto à competência técnica do instrutor

Os Observadores consideraram que os instrutores evidenciaram conhecer o conteúdo do Censo, foram capazes de dinamizar a aprendizagem dos treinandos e cumpriram plenamente as regras do treinamento, nas proporções apresentadas a seguir.

Gráfico 8.3 – Resultado da avaliação quanto à competência técnica do instrutor

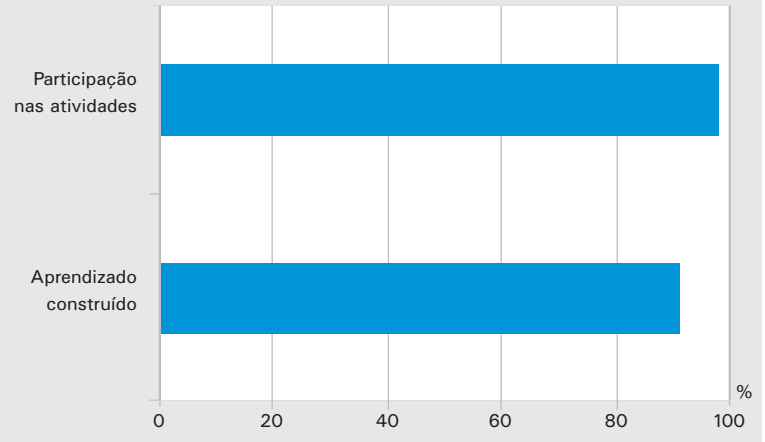


Fonte: Relatório de avaliação do programa de treinamento do censo demográfico 2000. Rio de Janeiro: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, 2000. 44 p.

c) quanto ao desempenho dos treinandos

Consideraram que os treinandos participaram ativamente das atividades didáticas e demonstraram ter assimilado corretamente o conteúdo ministrado as seguintes proporções de Observadores;

Gráfico 8.4 – Resultado da avaliação quanto ao desempenho dos treinandos

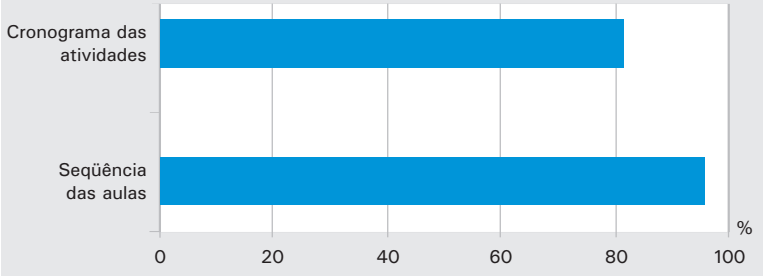


Fonte: Relatório de avaliação do programa de treinamento do censo demográfico 2000. Rio de Janeiro: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, 2000. 44 p.

d) sistemática do treinamento

Os Observadores foram de opinião de que a programação e o tempo de instrução previsto facilitaram a consecução dos objetivos do treinamento, e a seqüência proposta nos roteiros das aulas favoreceu a aprendizagem dos treinandos, nas seguintes proporções.

Gráfico 8.5 - Resultado da avaliação quanto à sistemática de treinamento



Fonte: Relatório de avaliação do programa de treinamento do censo demográfico 2000. Rio de Janeiro: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, 2000. 44 p.

8.4.2 Avaliação dos instrutores e treinandos

Essa atividade do treinamento do Censo foi executada através de uma avaliação de reação, elaborada por treinandos e instrutores. Teve como objetivo avaliar a qualidade do treinamento, levando em conta o desempenho dos instrutores, a adequação dos recursos didáticos, da carga horária e dos locais utilizados para a realização das aulas, entre outros aspectos.

A avaliação de reação consistiu de uma pesquisa realizada em uma amostra de instrutores e treinandos. Informações detalhadas sobre os critérios considerados no desenho da amostra podem ser obtidas no documento *Censo demográfico: pesquisa de avaliação dos treinamentos técnico-operacional e administrativo: aspectos de amostragem* (2000).

A seleção da amostra foi feita em dois estágios, dentro de cada unidade regional, segundo cada tipo de treinamento – técnico e administrativo/gerencial – e grupo de treinandos. No primeiro estágio, foram selecionadas as turmas e, em seguida, em cada uma delas, as unidades de segundo estágio, treinandos e instrutores. Alguns aspectos ligados ao método de seleção merecem citação:

- selecionada uma turma, todos os instrutores, automaticamente, estavam selecionados para a amostra;

- no Grupo 1, todos os instrutores e treinandos fizeram, necessariamente, parte da amostra da pesquisa;

- no Grupo 2, todas as turmas fizeram parte da amostra, sendo selecionados 10 treinandos em cada uma; e

- devido a características operacionais próprias do treinamento do Grupo 3, seus instrutores e treinandos não foram pesquisados e, portanto, não participaram da avaliação de reação.

As tabelas, apresentadas a seguir, fornecem informações sobre a amostra da pesquisa no Grupo 4 e no de Recenseadores. As diferenças entre os tamanhos das amostras, planejados e efetivos, ocorreram em função de problemas operacionais, que levaram a não aplicação da avaliação de reação, nas seguintes DIPEQs:

- PI, na fase de treinamento gerencial dos ACMs;

- RS, na fase de treinamento técnico do Grupo 4 e de Recenseadores; e

- TO, na fase de treinamento de Recenseadores.

Tabela 8.3 - Tamanho da amostra no treinamento técnico -2000

Grupos	Amostra planejada			Amostra efetiva		
	Turmas	Treinandos	Instrutores	Turmas	Treinandos	Instrutores
G4 + Recenseadores	676	2 028	1 352	583	1 749	1 146

Fonte: Censo demográfico 2000: pesquisa de avaliação dos treinamentos técnico-operacional e administrativo: aspectos de amostragem. Rio de Janeiro, IBGE, Diretoria de Pesquisas 2000. 67 p.

Tabela 8.4 - Tamanho da amostra no treinamento administrativo/gerencial - 2000

Grupos	Amostra planejada			Amostra efetiva		
	Turmas	Treinandos	Instrutores	Turmas	Treinandos	Instrutores
G4/ACM	191	573	382	163	489	332

Fonte: Censo demográfico 2000: pesquisa de avaliação dos treinamentos técnico-operacional e administrativo: aspectos de amostragem. Rio de Janeiro, IBGE, Diretoria de Pesquisas 2000. 67 p.

Cada instrutor ou treinando, selecionado para a amostra da pesquisa, fazia a avaliação de reação, atribuindo conceitos, que variavam de 1 – péssimo – a 5 – muito bom –, a diversos aspectos ligados ao treinamento, utilizando um dos seguintes questionários de avaliação:

Treinamento técnico para a coleta de dados

- TR 11 – Avaliação do Treinamento – Treinando
- TR 12 – Avaliação do Treinamento – Instrutor; e

Treinamento administrativo/gerencial

- TR 21 – Avaliação do Treinamento – Treinando
- TR 22 – Avaliação do Treinamento – Instrutor

Utilizando o TR 11 ou o TR 21, os treinandos avaliavam os seguintes aspectos: adequação dos recursos didáticos ao treinamento; tamanho da turma; carga horária total e sua distribuição diária; respeito aos horários; forma de apresentação e domínio do conteúdo pelo instrutor; relação entre os treinandos e destes com o instrutor; diversidade de atividades; e esclarecimento de dúvidas.

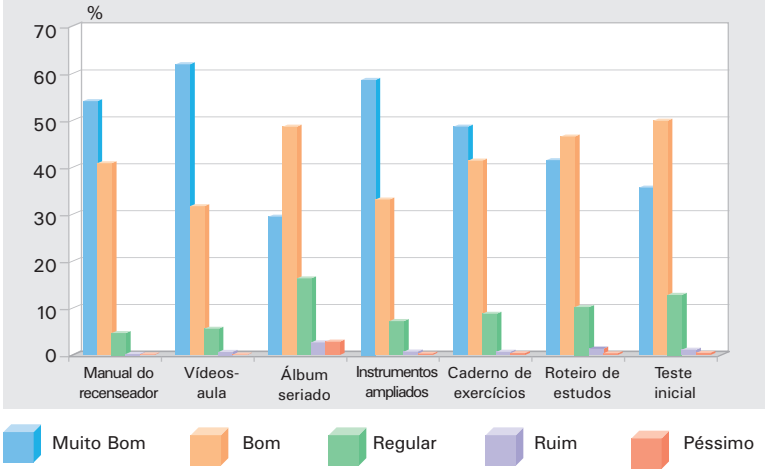
Já os instrutores, ao responderem ao TR 12 ou TR 22, faziam sua avaliação, opinando sobre a adequação dos seguintes pontos: recursos didáticos, relativamente ao treinamento e ao tempo previsto; roteiros de aula; atividade de auto-instrução – somente no TR 12; e alcance dos objetivos do treinamento.

O desenho da amostra da pesquisa foi elaborado tendo em vista obter estimativas com precisão estabelecida da proporção de atribuição de cada um dos conceitos de avaliação, em cada tipo de treinamento, para o Grupo 4, em nível de Brasil e, para o Grupo de Recenseadores, também para as cinco Grandes

Regiões do País. Essas estimativas tinham como referência cada um dos aspectos pesquisados nos questionários, sendo que a amostra também permitia resultados em separado, para instrutores e treinandos, em todos os níveis citados.

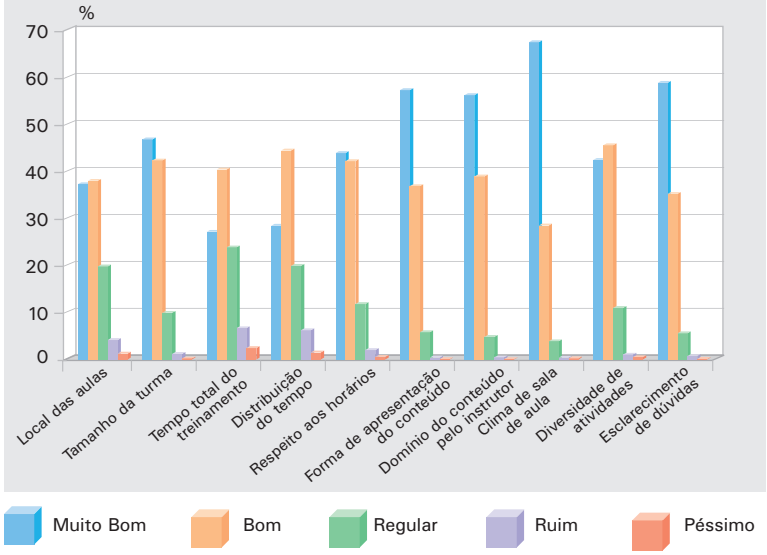
Os dois gráficos seguintes ilustram alguns resultados da avaliação dos instrutores e treinandos do Censo 2000.

Gráfico 8.6 - Avaliação dos instrutores
Adequação dos recursos didáticos ao treinamento



Fonte: Relatório de avaliação do programa de treinamento do censo demográfico 2000. Rio de Janeiro: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, 2000. 44 p.

Gráfico 8.7 - Avaliação dos treinandos



Fonte: Relatório de avaliação do programa de treinamento do censo demográfico 2000. Rio de Janeiro: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, 2000. 44 p.

Referências

CENSO demográfico 2000: instruções para revisão do trabalho da base operacional geográfica – TR 23. Rio de Janeiro: IBGE, 2000. 32 p.

_____: manual de coordenação técnica – CD 1.12. Rio de Janeiro: IBGE, 2000. 19 p.

_____: manual do agente censitário municipal – CD 1.11. Rio de Janeiro: IBGE, 2000. 33 p.

_____: manual do instrutor – TR 01. Rio de Janeiro: IBGE, 2000. 244 p.

_____. manual do instrutor – TR 18. Rio de Janeiro: IBGE, 2000. 72 p.

_____. manual do recenseador – CD 1.09. Rio de Janeiro: IBGE, 2000. 151 p.

_____. manual do sistema de administração de pessoal: orientações gerais – AD 01. Rio de Janeiro: IBGE, 2000. 45 p.

_____. manual do supervisor – CD 1.10. Rio de Janeiro: IBGE, 2000. 25 p.

_____. manual do usuário do sistema de indicadores gerenciais da coleta de dados – TR 24. Rio de Janeiro: IBGE, 2000. 37 p.

_____: pesquisa de avaliação dos treinamentos técnico operacional e administrativo: aspectos de amostragem. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2000. 67 p.

PROJETO de treinamento do Censo 2000. Rio de Janeiro: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, 1999. 44 p.

RELATÓRIO de avaliação do programa de treinamento do Censo Demográfico 2000. Rio de Janeiro: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos, 2000. 44 p.

9 Coleta das informações

A coleta das informações constitui-se numa das fases mais importantes de qualquer pesquisa. As principais preocupações com relação à coleta são: garantir a cobertura da população objetivo da pesquisa, especialmente importante no censo, que deve alcançar toda a população residente; assegurar a qualidade das informações levantadas; e atender ao cumprimento dos prazos estabelecidos.

Assim, em um censo, aspectos como o dimensionamento adequado das equipes de campo, o planejamento e execução cuidadosos das atividades de controle do trabalho e o suporte de um sistema de informações contribuem, decisivamente, para o sucesso da coleta.

Como grande parte das informações do censo é pesquisada por amostragem, é importante destacar, também, a necessidade da elaboração de um desenho amostral adequado, cujo sucesso dependerá, em grande medida, de uma execução rigorosa do trabalho de coleta dos dados.

9.1 Estratégias para a coleta

Além de recensear todos os moradores nos domicílios particulares ocupados e nos domicílios coletivos, o censo fez o cadastro ou listagem de todas as unidades residenciais – domicílios fechados, vagos ou de uso ocasional – bem como das unidades não-residenciais – lojas, escolas, etc. – que compunham cada setor, cujas informações constantes dos seus mapas e descrição dos limites foram também objeto de atualização.

Para dar suporte a todo esse trabalho, foram criados os postos de coleta, unidades de trabalho transitórias, estabelecidas em cada município, distrito, bairro ou localidade, para dar suporte à operação de coleta do censo. O posto de coleta constituiu-se na unidade mais descentralizada da estrutura censitária, onde se reunia a equipe encarregada da coleta de dados, em uma determinada área.

Os postos de coleta deviam ser instalados, preferencialmente, em regiões servidas por estruturas de transporte e comunicação que atendessem a padrões de segurança pessoal e patrimonial, bem como aos aspectos ligados ao sigilo da operação censitária. Deviam ter área compatível para abrigar as equipes e guardar todo o material – técnico, promocional, administrativo, de escritório e limpeza – necessário ao seu funcionamento.

Essas unidades podiam ser de dois tipos; postos informatizados e não-informatizados. Os primeiros, em número de 1000, dispunham de equipamentos de informática e linha telefônica para operação do Sistema de Indicadores Gerenciais da Coleta – SIGC e do Sistema de Administração de Pessoal – SAP, serviam de base para um Coordenador de Subárea e, preferencialmente, contavam com pessoal de apoio administrativo. Já os mais de 6000 postos não-informatizados, em suas necessidades administrativas e de comunicação, eram atendidos por postos informatizados ou por agências do IBGE.

Por outro lado, foi estabelecida uma estrutura de cargos de gerenciamento e operação dos trabalhos da coleta. Em cada unidade regional, existiam um Coordenador Técnico, Coordenadores de Área e de Subárea, Agentes Censitários Municipais, Agentes Censitários Supervisores e Recenseadores. Suas atribuições eram, em linhas gerais, as seguintes:

- Recenseador – técnico contratado para realizar a coleta das informações do censo em um setor;

- Agente Censitário Supervisor ou, simplesmente, Supervisor – técnico contratado para acompanhar, avaliar e, sobretudo, orientar uma equipe de, aproximadamente, oito Recenseadores;

- Agente Censitário Municipal – técnico contratado que teve a responsabilidade de todo o trabalho de instalação e organização do posto de coleta, que podia abranger um município ou parte dele, nos casos dos municípios de grande porte; durante a coleta das informações, desempenhou as funções de gerente de um grupo de Supervisores;

- Coordenador de Subárea – era o chefe da agência do IBGE, ou outro servidor da Instituição indicado para a função; em média, coordenou o trabalho de oito municípios;

- Coordenador de Área – foi o responsável por todo o trabalho de um conjunto de agências, em média, abrangendo três subáreas; e

- Coordenador Técnico – tinha a responsabilidade de coordenar todo o trabalho técnico da unidade regional.

A estratégia de controle e gerenciamento do trabalho de campo contou com o suporte do SIGC e do Sistema Auxiliar de Indicadores, que serão apresentados mais à frente.

Para o acompanhamento do trabalho de coleta de dados, em primeiro lugar, foi estabelecida uma rotina de supervisão e controle em âmbito local, que constou de um procedimento padrão de verificação da qualidade do material coletado, cuja aplicação esteve a cargo do Supervisor, e foi realizada durante todo o período da coleta de cada setor. Em complementação, existiram outros mecanismos de controle da coleta, que foram implementados pelos técnicos da Diretoria de Pesquisas: um realizado à distância, que resultava na elaboração de relatórios enviados aos Coordenadores Técnicos e mais o acompanhamento geral e o acompanhamento direcionado, baseados em visitas às unidades regionais.

9.2 Amostragem na coleta

Em 1960, o IBGE iniciou a utilização de amostragem na coleta dos dados dos censos demográficos. O procedimento adotado nos Censos 1960, 1970, 1980 e 1991 consistiu em aplicar um questionário básico a todos os domicílios e um segundo questionário, detalhado, a uma amostra de domicílios. Os domicílios foram selecionados para a amostra por meio de um procedimento sistemático com eqüiprobabilidade dentro de cada setor da base geográfica, de acordo com a fração amostral definida para cada Censo. Uma alternativa a esse procedimento é a utilização de amostragem de setores e a aplicação do questionário detalhado a todos os domicílios do setor selecionado (amostragem conglomerada de domicílios em um estágio).

O item seguinte contém o resumo das discussões a respeito dessas duas alternativas, realizadas durante o planejamento dos Censos anteriores, de 1980 e 1991. Logo em seguida, essa mesma questão é discutida com vistas ao planejamento do Censo Demográfico 2000, São apresentadas, ainda, as considerações finais e a alternativa recomendada e de fato utilizada.

9.2.1 Plano Amostral

O período de planejamento de um novo Censo é o momento oportuno para a avaliação dos métodos e procedimentos utilizados em todas as etapas de realização da pesquisa censitária, o que significa uma reavaliação também dos aspectos de amostragem. Assim, na etapa de planejamento do Censo Demográfico 2000, uma questão que novamente foi colocada em discussão é a que se refere ao plano amostral. A alternativa apresentada ao plano que vinha sendo adotado desde 1960 foi a mesma já considerada no planejamento dos dois últimos censos demográficos brasileiros, qual seja a adoção de plano amostral, onde o setor seria a unidade de seleção aleatória, e todos os domicílios pertencentes ao setor selecionado passariam a pertencer à amostra, ou seja, amostragem de conglomerados (setores) em um estágio.

Na realidade, o esquema de amostragem utilizado nos censos demográficos brasileiros não considera a amostragem aleatória simples de domicílios. Pelo fato de não se dispor de um cadastro atualizado de domicílios, a sua seleção é sistemática com eqüiprobabilidade, além de ser feita independentemente em cada setor; isto significa que cada setor é um estrato. Por um lado, quando há homogeneidade dentro dos conglomerados, um plano de amostragem de conglomerados em um estágio implica perda de precisão em relação à amostragem aleatória simples. Por outro lado, a estratificação de unidades homogêneas implica ganhos de precisão

também em relação à amostragem aleatória simples. Portanto, considerar cada setor como um estrato, com seleção de domicílios em todos os setores, torna-se mais eficiente do que o plano de amostragem aleatória simples de domicílios.

Silva e Moura (1990) realizaram estudos envolvendo estimativas para coeficientes de correlação intraclasse em nível de setor. Foi analisado um conjunto de 39 variáveis construídas com os dados do Censo Demográfico 1980. Os coeficientes de correlação intraclasse foram calculados com base no desenho amostral usado no Censo e foram obtidos para 9 regiões metropolitanas, além do Distrito Federal e do município de Goiânia. As conclusões do estudo apontam para a confirmação da hipótese de que o setor é unidade homogênea em relação à maioria das características de domicílio avaliadas.

Em função dos resultados apresentados em Silva e Moura (1990) e das considerações anteriores, concluiu-se que a adoção de amostragem de setores em um único estágio, com a inclusão de todos os domicílios dos setores selecionados para a amostra, levaria à obtenção de estimativas muito menos precisas do que as obtidas com uma amostra aleatória simples dos domicílios, de mesmo tamanho.

A adoção de um plano amostral por conglomerados em um estágio acarretaria ainda uma questão de extrema relevância, que diz respeito à limitação do nível geográfico de estimação, que precisaria ser definido *a priori*. Por outro lado, com a amostragem aleatória simples ou sistemática de domicílios em todos os setores, a área de estimação (ou área de ponderação) pode ser definida *a posteriori* e em níveis geográficos tão detalhados quanto se queira, com possibilidades de controlar a precisão das estimativas.

Por fim, a disponibilidade de informações da amostra, em todos os setores, facilita a realização de estudos para o planejamento das amostras das pesquisas domiciliares do período intercensitário, que tradicionalmente consideram o setor ou como unidade primária ou como unidade secundária de seleção do plano amostral. Desta forma, levando-se em conta: a busca por um plano amostral que resulte em melhor precisão; a demanda por informações em níveis geográficos pequenos; e a utilidade de informações da amostra em todos os setores, para o planejamento de amostras no período intercensitário, decidiu-se pela manutenção do plano amostral tradicional, adotado na coleta dos censos demográficos anteriores. Assim, tem-se a seleção sistemática de domicílios particulares e de famílias e pessoas sós em domicílios coletivos, feita independentemente em cada setor, o que equivale a um plano de amostragem estratificada, onde os estratos são os setores, e a seleção é sistemática com eqüiprobabilidade em cada estrato.

9.2.2 Tamanho da amostra

Nos Censos de 1960, 1970 e 1980, a fração amostral adotada foi de 25%, ou seja, um a cada quatro domicílios. No Censo Demográfico 1991, foram adotadas duas frações amostrais distintas: 20% para os municípios com população estimada até 15 000 habitantes e 10% para os demais municípios.

Para o Censo 2000, também estava prevista a utilização de dois questionários, um deles a ser aplicado em uma amostra de domicílios. O plano amostral a ser utilizado já havia sido definido como o mesmo adotado nos censos precedentes, conforme descrito no item anterior, restando definir a fração amostral a ser usada em cada município. Assim, uma avaliação das discussões e estudos realizados quando do planejamento dos censos passados, bem como o

estabelecimento dos objetivos (ou das premissas) para o Censo 2000, formaram a base para a definição da fração amostral que seria definida para a investigação do questionário mais detalhado no Censo 2000.

Planejamento do Censo Demográfico 2000

Uma questão que novamente se colocou em discussão, durante a fase de planejamento, é a que se refere à fração amostral.

É fato comprovado que questionários muito grandes aumentam os custos e prazos de uma pesquisa, aumento esse que pode ser apenas parcialmente compensado pela diminuição do tamanho da amostra. Além disso, questionários extensos também aumentam a incidência dos chamados erros alheios à amostragem, que podem ser mais controlados com uma amostra de tamanho menor, desde que tomadas as devidas medidas.

A tabela 9.1, a seguir, apresenta o número de municípios existentes classificados por algumas classes de tamanho em população. A população considerada é aquela estimada pelo Departamento de População e Indicadores Sociais - DEPIS, da Diretoria de Pesquisas – DPE, do IBGE, e publicada no Diário Oficial da União em 31 de agosto de 1997, atendendo artigo 102 da Lei nº 8.443, de 16 de julho de 1992.

Tabela 9.1 - Número de municípios existentes em 1997, por classes de tamanho em população estimada para 1997

Classes de tamanho em população estimada	Número de municípios			
	Freqüência absoluta		Freqüência relativa (%)	
	Simples	Acumulada	Simples	Acumulada
Total	**5 507**	-	**100,0**	-
Até 5 000	1 408	1 408	25,6	25,6
De 5 001 a 10 000	1 303	2 711	23,7	49,2
De 10 001 a 15 000	865	3 576	15,7	64,9
De 15 001 a 30 000	1 068	4 644	19,4	84,3
De 30 001 a 50 000	377	5 021	6,8	91,2
De 50 001 a 100 000	280	5 301	5,1	96,3
De 100 001 a 250 000	134	5 435	2,4	98,7
De 250 001 a 500 000	44	5 479	0,8	99,5
De 500 001 ou mais	28	5 507	0,5	100,0

Fonte: IBGE, Diretoria de Geociências, Departamento de Estruturas Territoriais.

Se fossem definidas para o Censo 2000 as mesmas frações amostrais e arbitrados como ponto de corte os mesmos 15 000 habitantes no município, a situação em termos de número de municípios por fração amostral seria a seguinte.

Tabela 9.2 - Número de municípios existentes em 1997, por classes de tamanho em população estimada para 1997

Classes de tamanho em população estimada	Número de municípios	
	Freqüência absoluta	Freqüência relativa (%)
Total	**5 507**	**100,0**
Até 15 000 hab. (f = 20%)	3 576	64,9
Mais de 15 000 hab. (f = 10%)	1 931	35,1

Fonte: IBGE, Diretoria de Geociência, Departamento de Estruturas Territoriais.

Albieri (1998) reúne as questões referentes à fração amostral usada na coleta dos dados dos censos demográficos do Brasil e revisa as discussões ocorridas por ocasião do planejamento dos Censos de 1980 e 1991.

Foram realizados novos estudos com base nos dados do Censo Demográfico 1991, levando-se em conta frações alternativas, de acordo com a população dos municípios. É sob essa ótica que Bianchini e Albieri (1999b) apresentam um estudo de alternativas para as frações de amostragem do Censo Demográfico 2000, visando à avaliação dos efeitos da alteração das frações amostrais sobre a precisão das estimativas, bem como a avaliação da redução de custos e de tempo, associadas com os aspectos operacionais da coleta.

Ao todo, foram analisadas seis propostas de composição das frações amostrais, de acordo com os cortes populacionais definidos, considerando-se a população dos municípios projetada para o ano 2000 e o número de municípios existentes em 1997. A Comissão Consultiva do Censo Demográfico 2000 recomendou a manutenção das frações adotadas em 1991 e sugeriu a avaliação do aumento da fração de 20% para 50% nos municípios menos populosos. No entanto, restrições orçamentárias levaram o IBGE para outra direção: calcular os efeitos da redução da fração de 10% para 5%, nos municípios com mais de 100 000 habitantes.

As tabelas 9.3 e 9.4 apresentam, respectivamente, o número esperado de domicílios e de pessoas com Questionários da Amostra e Básico para o Censo Demográfico 2000, de acordo com os cortes populacionais definidos, considerando-se a população dos municípios projetada para o ano 2000 e o número de municípios existentes em 2000. Nessas tabelas, não estão incluídas as informações relativas a duas alternativas também estudadas, e que foram descartadas logo no início dos estudos, a saber: a utilização de fração amostral única de 10% para todos os municípios e a redução pela metade da fração amostral usada no Censo de 1991, ou seja, 10% para os municípios com até 15 000 habitantes e 5% para os demais municípios. Essas alternativas foram descartadas por acarretarem perda de precisão, de forma acentuada, nos municípios menos populosos.

Assim, verificou-se que tal redução não causaria perda expressiva de precisão para as estimativas no nível de município. Comparado com as frações usadas em 1991, o novo critério acarretaria uma redução de 23,3% de domicílios e de 15,1% de pessoas a serem investigadas na amostra, bem como um aumento de 3% de domicílios e de 2% de pessoas a serem investigadas com o Questionário Básico. A principal desvantagem seria a utilização de três frações: 20% para os municípios com menos de 15 000 habitantes, 10% para os que possuem mais de 15 000 e menos de 100 000 habitantes, e 5% para os municípios com população projetada igual ou maior do que 100 000 habitantes. Isto exigiria maior controle da operação de coleta e alguma perda de detalhamento geográfico das estimativas para os municípios maiores. Isto posto, a decisão final foi a de manter as mesmas frações usadas em 1991: 20% para os municípios com até 15 000 habitantes e 10% para os municípios com mais de 15 000 habitantes, o que corresponde à primeira alternativa apresentada nas tabelas 9.3 e 9.4.

Tabela 9.3 - Número esperado de domicílios investigados com questionários de amostra e básico para o Censo Demográfico, considerando-se frações amostrais alternativas, de acordo com a população dos municípios - 2000

Frações amostrais alternativas	Domicílios		
	Total	Questionário da amostra	Questionário básico
Mesmas frações de 91 (20% e 10%)	**42 508 736**	**4 842 755**	**37 665 981**
Municípios com até 15000 hab. (f = 20%)	5 918 817	1 183 763	4 735 054
Municípios com mais de 15000 hab. (f = 10%)	36 589 919	3 658 992	32 930 927
3 frações (20%, 10% e 5%)	**42 508 736**	**3 714 940**	**38 793 796**
Municípios com até 15000 hab. (f = 20%)	5 918 817	1 183 763	4 735 054
Municípios com 15001 a 100000 hab. (f = 10%)	14 033 618	1 403 362	12 630 256
Municípios com mais de 100000 hab. (f = 5%)	22 556 301	1 127 815	21 428 486
Variação (frações 20%, 10% e 5% / frações de 91)			
Absoluta	-	1 127 815	1 127 815
Relativa (%)	0,0	(-) 23,3	3,0
2 frações (50% e 10%)	**42 508 736**	**6 618 401**	**35 890 336**
Municípios com até 15000 hab. (f = 50%)	5 918 817	2 959 409	2 959 409
Municípios com mais de 15000 hab. (f = 10%)	36 589 919	3 658 992	32 930 927
Variação (frações 50% e 10% / frações de 91)			
Absoluta	-	1 775 645	(-) 1 775 645
Relativa (%)	0,0	36,7	(-) 4,7
3 frações (50%, 10% e 5%)	**42 508 736**	**5 490 585**	**37 018 151**
Municípios com até 15000 hab. (f = 50%)	5 918 817	2 959 409	2 959 409
Municípios com 15001 a 100000 hab. (f = 10%)	14 033 618	1 403 362	12 630 256
Municípios com mais de 100000 hab. (f = 5%)	22 556 301	1 127 815	21 428 486
Variação (frações 50%, 10% e 5% / frações de 91)			
Absoluta	-	647 830	(-) 647 830
Relativa (%)	0,0	13,4	(-) 1,7

Fonte: IBGE, Diretoria de Pesquisas, Departamento de População e Indicadores Sociais

Tabela 9.4 - Número esperado de pessoas investigadas com questionários de amostra e básico para o Censo Demográfico, considerando-se frações amostrais alternativas, de acordo com a população dos municípios - 2000

Frações amostrais alternativas	Pessoas		
	Total	Questionário da amostra	Questionário básico
Mesmas frações de 91 (20% e 10%)	**166 112 518**	**19 028 425**	**147 084 093**
Municípios com até 15000 hab. (f = 20%)	24 171 734	4 834 347	19 337 387
Municípios com mais de 15000 hab. (f = 10%)	141 940 784	14 194 078	127 746 706
3 frações (20%, 10% e 5%)	**166 112 518**	**16 161 145**	**149 951 373**
Municípios com até 15000 hab. (f = 20%)	24 171 734	4 834 347	19 337 387
Municípios com 15001 a 100000 hab. (f = 10%)	84 595 176	8 459 518	76 135 658
Municípios com mais de 100000 hab. (f = 5%)	57 345 608	2 867 280	54 478 328
Variação (frações 20%, 10% e 5% / frações de 91)			
Absoluta	-	2 867 280	2 867 280
Relativa (%)	0,0	(-) 15,1	2,0
2 frações (50% e 10%)	**166 112 518**	**26 279 945**	**139 832 573**
Municípios com até 15000 hab. (f = 50%)	24 171 734	12 085 867	12 085 867
Municípios com mais de 15000 hab. (f = 10%)	141 940 784	14 194 078	127 746 706
Variação (frações 50% e 10% / frações de 91)			
Absoluta	-	7 251 520	7 251 520
Relativa (%)	0,0	38,1	(-) 4,9
3 frações (50%, 10% e 5%)	**166 112 518**	**23 412 665**	**142 699 853**
Municípios com até 15000 hab. (f = 50%)	24 171 734	12 085 867	12 085 867
Municípios com 15001 a 100000 hab. (f = 10%)	84 595 176	8 459 518	76 135 658
Municípios com mais de 100000 hab. (f = 5%)	57 345 608	2 867 280	54 478 328
Variação (frações 50%, 10% e 5% / frações de 91)			
Absoluta	-	4 384 240	4 384 240
Relativa (%)	0,0	23,0	(-) 3,0

Fonte: IBGE, Diretoria de Pesquisas, Departamento de População e Indicadores Sociais

A tabela 9.5 apresenta o número de municípios por fração amostral planejada, de acordo com o tamanho medido pela população projetada para o ano 2000. A tabela 9.6 informa também o número de domicílios e de pessoas com Questionários da Amostra e Básico esperados.

Tabela 9.5 - Número de municípios existentes, segundo as classes de tamanho da população e a fração amostral planejada - 2000

Classes de tamanho da população e fração amostral planejada	Número de municípios	
	Freqüência absoluta	Freqüência relativa (%)
Total	**5 507**	**100,0**
Até 15 000 hab. (f = 20%)	3 533	64,2
Mais de 15 000 hab. (f = 10%)	1 974	35,8

Fonte: IBGE, Diretoria de Pesquisas, Departamento de População e Indicadores Sociais.

Tabela 9.6 - Número esperado de domicílios e de pessoas, por fração amostral planejada, segundo as unidades de investigação e o tipo de questionário

Unidades de investigação e tipo de questionário	Número esperado		
	Total	Fração amostral	
		20%	10%
Domicílios	**42 508 736**	**5 918 817**	**36 589 919**
Questionário da amostra	4 842 755	1 183 763	3 658 992
Questionário básico	37 665 981	4 735 054	32 930 927
Pessoas	**166 112 518**	**24 171 734**	**141 940 784**
Questionário da amostra	19 028 425	4 834 347	14 194 078
Questionário básico	147 084 093	19 337 387	127 746 706

Fonte: IBGE, Diretoria de Pesquisas, Departamento de População e Indicadores Sociais.

9.2.3 Seleção da amostra

O princípio fundamental da amostragem probabilística é o de ter as suas unidades selecionadas aleatoriamente. A seleção das unidades investigadas por amostragem constitui uma fase importante da coleta do censo, pois, ao longo da mesma podem ocorrer sérios inconvenientes nas informações coletadas, conduzindo a diversos tipos de tendenciosidades. Utilizando-se o domicílio como unidade de amostragem, e sendo realizada a seleção pelos Recenseadores, durante a coleta, as regras de seleção devem ser simples e de fácil aplicação. Por esse motivo, são adotados processos sistemáticos de seleção, estabelecidos de forma a quebrar ou interromper seqüências cíclicas do universo, obtendo-se, assim, bastante variação nas séries sistemáticas de seleção, ao mesmo tempo que se evita o surgimento de tendenciosidades cometidas durante a coleta.

A amostragem sistemática prescinde da existência prévia de cadastro e permite que ele seja construído simultaneamente à seleção da amostra, de tal forma que a operacionalização do processo de seleção (por intervalos prefixados), dentro de cada setor, é simples e de fácil aplicação pelos Recenseadores.

O cadastro para a seleção da amostra de domicílios particulares de cada setor é construído a partir do registro nas Folhas de Coleta de cada unidade visitada e identificada, conforme a sua espécie: domicílio particular ocupado,

domicílio particular fechado, domicílio particular de uso ocasional, domicílio particular vago, domicílio coletivo e não-residencial. Nas Folhas de Coleta de cada setor, são numerados, em ordem crescente, todos os domicílios particulares relacionados, quer sejam ocupados, fechados, de uso ocasional ou vagos, constituindo o cadastro para a seleção dos domicílios particulares. Para cada setor com domicílio coletivo registrado nas Folhas de Coleta, são preparadas Folhas de Domicílio Coletivo, com o registro e numeração das famílias e pessoas sós moradores, constituindo o cadastro para a seleção em domicílios coletivos, independente do cadastro de seleção dos domicílios particulares. Este processo de trabalho está descrito no item 9.4, deste mesmo capítulo, onde estão descritos os trabalhos sob responsabilidade do Recenseador.

Processo de seleção

A seleção da amostra foi feita de acordo com a fração amostral definida, sendo que em cada município foi usada somente uma fração, 10% ou 20%. Os domicílios particulares foram selecionados com base no Número de Ordem (Coluna 5 da Folha de Coleta), e o processo de seleção foi feito com base nos seguintes procedimentos:

Para os municípios onde foi adotada a fração amostral de 10%, foram selecionados para a amostra os domicílios particulares, cuja unidade do Número de Ordem fosse igual à unidade do número do setor, somando-se um intervalo de dez unidades. Assim, em um setor de número 0012, foram selecionados os domicílios com os números de ordem na Folha de Coleta 2, 12, 22, ..., enquanto em um setor de número 0010 foram selecionados os domicílios com os números de ordem 10, 20, 30, e assim sucessivamente.

Para os municípios onde foi adotada a fração amostral de 20%, foram selecionados para amostra os domicílios particulares da Folha de Coleta, tomando como base a unidade do número do setor, usando-se um intervalo de cinco unidades, conforme o Quadro 9.1 a seguir.

Quadro 9.1 - Indicação das unidades a serem selecionadas no setor censitário, de acordo com o número do setor

Setores terminados em: (unidade do número do setor)	Primeira unidade selecionada para a amostra	Unidades selecionadas no setor
1 ou 6	1	1, 6, 11, 16, 21, 26...
2 ou 7	2	2, 7, 12, 17, 22, 27...
3 ou 8	3	3, 8, 13, 18, 23, 28...
4 ou 9	4	4, 9, 14, 19, 24, 29...
5 ou 0	5	5, 10, 15, 20, 25, 30...

Assim, em um setor de número 0027, foram selecionados os domicílios com os seguintes números de ordem na Folha de Coleta: 2, 7, 12, 17, 22, ..., enquanto em um setor de número 0013 foram selecionados os domicílios com os números de ordem 3, 8, 13, 18, e assim sucessivamente. Em um setor de número 0128, foram selecionados os domicílios com os mesmos números de ordem, na Folha de Coleta, que os do exemplo anterior, o setor de número 0013. Este mesmo processo foi utilizado na Folha de Domicílio Coletivo, para a

seleção das famílias e pessoas sós ali moradoras, com base na numeração da coluna 1, Número de Ordem. Desta forma, a amostra foi formada por domicílios particulares e pessoas neles residentes, bem como pelas famílias ou pessoas sós moradoras em domicílios coletivos.

No caso de ter sido selecionado um domicílio particular fechado, de uso ocasional ou vago, não houve preenchimento do Questionário da Amostra, mas foi mantida a série da seleção.

9.3 Instrumentos de coleta

Os instrumentos de coleta utilizados no Censo Demográfico 2000 foram sempre objeto de atenção de especialistas, tendo como objetivo torná-los o mais eficientes possível. Além dos cuidados para definição de seus conteúdos, sua forma final procurou sempre atender aos aspectos de apresentação e, sobretudo, aos diversos requisitos técnicos exigidos pelo trabalho de campo. Assim, no caso dos textos dos manuais de instrução, para sua perfeita compreensão, foram adotadas uma linguagem direta e uma forma de apresentação agradável à leitura, por meio de recursos gráficos de realce, utilização de figuras, etc. Para facilitar a procura dos assuntos, utilizaram-se ícones de identificação, além do índice. No que toca aos questionários e outros formulários para registro das informações, teve-se, inicialmente, a preocupação quanto ao seu manuseio e portabilidade no campo – dimensões adequadas e qualidade do papel - além, é claro, de se buscar uma apresentação – disposição dos quesitos, indicadores gráficos, comandos de seqüência, etc. – que reduzisse a possibilidade de erros no preenchimento.

Por fim, é importante citar que cuidados especiais de impressão foram dispensados aos instrumentos de coleta, para cujos registros seriam utilizados procedimentos de captura, através da digitalização de imagens e do reconhecimento ótico de caracteres. Para maiores detalhes a respeito, consultar o capítulo 3, item 3.2.1.

A seguir, são identificados os diferentes instrumentos de coleta, acompanhados por uma descrição de suas características. Todos estão apresentados no anexo de CD-ROM; já os CDs 1.01, 1.02, 1.03, 1.06 e 1.07 também constam do anexo deste capítulo.

CD 1.01 - Questionário Básico

Instrumento de coleta, em forma de caderno, utilizado para o registro das características do domicílio e dos seus moradores, na data de referência, em cada unidade domiciliar ocupada que não foi selecionada para a amostra. Contém quatro páginas, podendo registrar informações para até seis pessoas moradoras. Na primeira página, além do nome do município e da agência do IBGE, constam, no bloco 1, os itens de identificação e, no bloco 2, os quesitos de caracterização do domicílio. Na segunda página, tem-se, no bloco 3, a lista de moradores do domicílio e, iniciando o bloco 4, Características dos Moradores, as informações referentes à pessoa responsável ou o morador individual em domicílio coletivo. Por fim, as páginas 3 e 4 davam continuidade ao bloco 4, apresentando os quesitos destinados aos demais moradores, da segunda até a sexta pessoa.

CD 1.02 - Questionário da Amostra

Instrumento de coleta mais detalhado, em forma de caderno, utilizado para o registro das características do domicílio e dos seus moradores, na data de referência, em cada unidade domiciliar ocupada que foi selecionada para a amostra. Contém trinta e duas páginas, podendo, a exemplo do CD 1.01, serem registradas informações para até seis moradores. Na primeira página existe espaços para registro dos nomes do município e da agência do IBGE, para os itens de identificação – bloco 1 – e para os dez primeiros quesitos do bloco 2, Características do Domicílio. Na segunda página, constam os demais quesitos do bloco 2 e o bloco 3, que corresponde à lista de moradores do domicílio. Na página 3, tem-se o início do bloco 4, Características do Morador, que abrange os quesitos relativos a sexo, idade, relação com a pessoa responsável pelo domicílio e pela família, idade, cor ou raça, religião, deficiência mental permanente e auto avaliação da capacidade física. A página 4 é ocupada integralmente pelos quesitos relativos à migração e deslocamento, enquanto na página seguinte são encontrados os quesitos sobre educação e os que investigam os estados civil e conjugal, para os moradores com dez anos ou mais de idade, na data de referência.

Na página 6, tinha-se o início do levantamento das características sobre trabalho e rendimento, para os moradores com 10 anos ou mais de idade, complementado na página 7, que apresentava também, para as mulheres de 10 anos ou mais de idade, os quesitos sobre fecundidade e mortalidade, encerrando a investigação sobre a primeira pessoa. Nas páginas seguintes, de números 8 a 32, repetiam-se os quesitos do bloco 4 – Características do Morador, para a segunda pessoa, até se chegar à sexta do domicílio.

CD 1.03 - Folha de Domicílio Coletivo

Utilizada para listar as famílias e pessoas sós que foram recenseadas em cada domicílio coletivo, além de servir para a seleção das unidades nas quais seria aplicado o Questionário da Amostra.

CD 1.06 - Caderneta do Setor

Instrumento cujo preenchimento marcava o início e o encerramento da atividade de coleta no setor. Apresenta o mapa e a descrição do setor, destinando-se, basicamente, ao registro do resumo das informações coletadas – total de unidades visitadas, de questionários aplicados e de pessoas recenseadas. Serve também para caracterizar o setor quanto à existência de identificação, iluminação pública e calçamento/pavimentação, bem como para guardar as Folhas de Coleta e as Folhas de Domicílio Coletivo utilizadas.

CD 1.07 - Folha de Coleta

Utilizada para o registro das unidades residenciais e não-residenciais existentes no setor e do número de moradores e de questionários utilizados em cada domicílio ocupado, servindo também para a seleção dos domicílios particulares nos quais seria aplicado o Questionário da Amostra. A caracterização do trecho do logradouro percorrido pelo Recenseador, quanto à existência de identificação, iluminação pública e calçamento/pavimentação, também era feita nesse instrumento.

CD 1.09 - Manual do Recenseador

Instrumento que deu suporte para todo o trabalho do Recenseador, sendo utilizado desde a etapa de seu treinamento. Nele, constam os conceitos, as definições, os procedimentos e as orientações necessárias ao desempenho de sua tarefa, bem como as normas de preenchimento dos instrumentos de coleta.

CD 1.10 - Manual do Supervisor

Contém todas as instruções que orientam o trabalho do Supervisor, desde o conhecimento e a distribuição dos setores aos Recenseadores até o encerramento da coleta e o preparo do material para remessa às agências do IBGE. Instrui, detalhadamente, sobre todos os procedimentos para avaliação do material coletado, com regras específicas para realização de novas entrevistas e ações gerenciais daí decorrentes.

CD 1.11 - Manual do Agente Censitário Municipal – ACM

Instrumento voltado para as tarefas de responsabilidade do ACM, com ênfase especial nos procedimentos operacionais do SIGC e nas ações gerenciais a serem realizadas junto aos Supervisores e às instâncias superiores da coordenação técnica estadual do censo.

CD 1.12 - Manual de Coordenação Técnica

Descreve as tarefas a cargo da coordenação estadual, tendo em conta sua responsabilidade com relação à condução do censo e ao vínculo com a equipe técnica central. Assim, nesse manual foram reunidas todas as instruções e procedimentos necessários ao desempenho das atribuições, destacando-se os aspectos gerenciais ligados ao SIGC, inclusive no que diz respeito à autorização de pagamento dos Recenseadores, bem como à guarda e preservação do material de coleta.

CD 1.13 - Guia de Remessa dos Questionários

Documento utilizado pelo ACM, nos postos não-informatizados, para remessa das pastas ao Coordenador de Subárea, contendo os questionários CD 1.01 e CD 1.02 de cada setor, que teve o trabalho de coleta encerrado e estava liberado pelo SIGC.

CD 1.14 - Guia de Remessa das Folhas de Coleta

Documento utilizado pelo ACM, após o término dos trabalhos do posto de coleta não-informatizado, para remessa ao Coordenador de Subárea das pastas com os CD 1.07 de cada setor.

CD 1.15 – Agenda do Supervisor

Tinha por objetivo manter, de forma organizada, um histórico dos trabalhos de coleta do pessoal sob sua responsabilidade. Servia para controlar o esquema de visitas dos Recenseadores ao posto de coleta e para registrar as informações relativas à coleta semanal dos setores.

CD 1.16 – Registro dos Setores Iniciados

Documento preenchido semanalmente pelo Supervisor, indicando ao ACM a data efetiva do início da coleta em cada setor, bem como a matrícula do Recenseador responsável pela sua coleta.

CD 1.17 – Resumo da Produção Semanal

Documento preenchido pelo Supervisor, que registrava a produção semanal acumulada do Recenseador – total de questionários aplicados e de pessoas recenseadas – no setor sob sua responsabilidade. Esse resumo servia também, para que o ACM fizesse a carga dos dados no SIGC, permitindo, assim, o acompanhamento gerencial do avanço do trabalho de coleta do Censo.

TR 23 – Instruções para Revisão do Trabalho da Base Operacional Geográfica

Manual que orientou todo o trabalho do Supervisor, antes do início da visita aos domicílios. Tinha por objetivo rever o trabalho da Base Operacional Geográfica – BOG, que incluía também a correção do Cadastro de Segmento de Logradouros – CADLOG, nos setores urbanos ou de aglomerados rurais, dos municípios com população, dada pela Contagem da População de 1996, acima de 25 000 habitantes. Além de reunir todos os conceitos necessários à realização da tarefa, informa, detalhadamente, sobre a evolução da atividade, desde a avaliação inicial no escritório, passando pelo reconhecimento em campo da área de trabalho, chegando até as correções e complementações nos mapas.

9.4 Procedimentos de atuação no campo

Serão apresentadas aqui, de maneira breve, as principais atividades desenvolvidas pela equipe de campo do censo, começando pelas de responsabilidade do Recenseador, com destaque para os procedimentos de percurso e cobertura do setor, bem como as recomendações para a boa condução das entrevistas. A seguir, serão tratadas as atividades de supervisão local, com ênfase especial para a crítica do material coletado, bem como as tarefas do Agente Censitário Municipal – ACM – técnico responsável pelo posto de coleta. Por fim, serão descritas as atribuições ligadas aos Coordenadores Técnicos, ressaltando o aspecto gerencial de seu trabalho.

9.4.1 Trabalho do Recenseador

No censo, uma perfeita cobertura do setor acontece quando todas as suas unidades, residenciais ou não-residenciais, são listadas de acordo com as regras estabelecidas, e todos os seus domicílios são visitados para realização das entrevistas. Assim, o Recenseador deve estar atento para não incorrer no erro de evasão, que subenumera unidades e pessoas residentes. Da mesma forma, não deve incluí-las indevidamente, tendo o cuidado, ao percorrer o setor, de não invadir setores vizinhos, fazendo com que unidades sejam listadas mais de uma vez, acarretando a superenumeração da população.

O Recenseador, de posse do CD 1.06, antes de dar início às entrevistas nas áreas urbanas, teve como incumbência conhecer o seu setor de trabalho, percorrendo-o integralmente, identificando os seus limites, logradouros e os

pontos, inicial e final, do percurso a ser cumprido. Todas as dúvidas levantadas deveriam ser apresentadas ao Supervisor, o mais rápido possível, para não comprometer os prazos estabelecidos.

Para mostrar como são percorridos os setores, é importante distinguí-los da seguinte maneira: setores urbanos, ou aglomerados rurais com quarteirões fechados ou abertos; setores urbanos ou aglomerados rurais não-divididos em quarteirões; e setores rurais.

Para os setores urbanos ou aglomerados rurais, com quarteirões fechados ou abertos, o caminho do Recenseador começava no ponto inicial marcado no mapa, percorrendo o setor, um quarteirão de cada vez, mantendo a área de trabalho sempre à sua direita, até alcançar o ponto final.

No caso dos setores urbanos ou aglomerados rurais não-divididos em quarteirões, o percurso do Recenseador começava no ponto inicial, continuando rua por rua, estrada por estrada, percorrendo um de seus lados de cada vez, mantendo sempre a área de trabalho à sua direita, até o ponto final do setor.

Nos dois tipos de setores citados, quando houvesse logradouros transversais (ruas, vielas, becos e caminhos), o Recenseador interrompia o percurso da via principal e percorria integralmente esses logradouros, retornando em seguida ao ponto onde foi feita a interrupção, continuando, a partir daí, o seu trabalho.

Quando se tratava de uma área rural, não havia regras fixas para que o setor fosse percorrido. O princípio do percurso, a exemplo dos demais tipos de setores, era feito a partir do ponto inicial. No entanto, caso esse ponto marcado no mapa fosse de difícil acesso, o Recenseador estabelecia um outro, desde que situado nos limites do setor. A partir daí, o caminho a ser percorrido no setor era determinado pela localização de suas unidades, ou por informações dadas pelos moradores, indicando qual a casa mais próxima e a forma mais fácil de se chegar a ela. Nessa última situação, era necessária atenção redobrada para evitar a invasão de setores vizinhos.

O Recenseador, durante o percurso, fazia a listagem do setor, ou seja, registrava nas Folhas de Coleta – CD 1.07 – todas as unidades residenciais e não-residenciais e, para cada domicílio coletivo encontrado, relacionava as famílias ou as pessoas sós, nas Folhas de Domicílio Coletivo – CD 1.03.

Ao mesmo tempo que as unidades eram listadas, fazia-se a seleção da amostra do censo, indicando ao Recenseador em quais unidades seria aplicado o CD 1.01 - Questionário Básico – ou o CD 1.02 – Questionário da Amostra. O processo de seleção da amostra foi feito independentemente para os domicílios particulares e para as famílias ou pessoas sós, nos domicílios coletivos.

À medida que o Recenseador fazia os lançamentos, na ordem do percurso, em cada linha da Folha de Coleta, classificava, na coluna 2.03, as unidades segundo a espécie. As identificadas com códigos um a quatro, respectivamente, domicílios particulares ocupados, fechados, de uso ocasional ou vagos, recebiam numeração seqüencial na coluna 2.05 – Número de Ordem Domicílio Particular . Para o primeiro domicílio coletivo encontrado, o Recenseador listava as famílias ou pessoas sós, em cada linha do CD 1.03, e as numerava seqüencialmente na coluna 1 – Número de Ordem. Se, no percurso do setor, um segundo domicílio coletivo fosse encontrado, uma nova Folha de Domicílio Coletivo era aberta, mas a numeração de suas unidades começava a partir do número dado à última família ou pessoa só listada do domicílio coletivo anterior, e assim sucessivamente.

À medida que as duas numerações avançavam, de um lado, a dos domicílios particulares e, de outro, a das famílias ou pessoas sós nos domicílios coletivos, eram selecionadas as unidades da amostra, respeitando-se as frações de amostragem – 10% ou 20% – previamente estabelecidas para cada município. A rotina de seleção era seqüencial, e seu início levava em conta o algarismo das unidades do Número do Setor – campo 1.06 da Folha de Coleta. Para maiores esclarecimentos, consultar o item 9.2.3.

Quando da visita às unidades, o Recenseador, estava obrigado a se apresentar portando o crachá de identificação. Durante a entrevista, devia evitar assuntos estranhos ao trabalho de coleta. As perguntas deviam ser feitas de forma direta, como estavam redigidas nos questionários. Somente depois disso, se necessário, esclarecimentos adicionais eram feitos. Especial atenção era dispensada à elaboração da lista de moradores, buscando evitar a omissão de pessoas que, no momento da entrevista, não estivessem presentes, ou de crianças recém-nascidas. Do mesmo modo, a observância aos conceitos, definições e formas de investigação, contidas no Manual do Recenseador, era imprescindível. Encerrada a entrevista, o agradecimento ao informante era obrigatório, além do aviso de que outro técnico do IBGE poderia retornar para esclarecimento de alguma questão.

Quando em uma unidade domiciliar, seus moradores estivessem temporariamente ausentes, o Recenseador devia, obrigatoriamente, retornar para obtenção das informações, de modo que a condição de domicílio fechado somente fosse caracterizada após esgotados todos os esforços para obtenção da entrevista.

Quando ocorresse recusa na prestação de informações, o Recenseador tentava, amistosamente, convencer a pessoa a conceder a entrevista. Não obtendo sucesso, comunicava o fato, imediatamente, a seu Supervisor.

Para o bom andamento do trabalho, o Recenseador comparecia ao posto de coleta, no mínimo, uma vez por semana, na data agendada pelo seu Supervisor, para a entrega do material coletado e esclarecimento de dúvidas.

9.4.2 Trabalho do Agente Censitário Supervisor - ACS

No censo, o Supervisor desempenhava atividades técnicas e administrativas. Assim, antes da chegada dos Recenseadores ao posto de coleta, verificava se todo o material a ser utilizado estava disponível na quantidade necessária ao trabalho de campo a cargo de sua equipe.

Da mesma forma, antes do início da coleta, utilizando o CD 1.06, o Supervisor, fazia o reconhecimento das áreas de trabalho sob sua responsabilidade, revisando os mapas dos setores e as descrições dos seus limites, numerando, se houvesse, seus quarteirões e as respectivas faces. Nos municípios cuja população superou 25 000 habitantes na Contagem da População 1996, realizava também a revisão do Cadastro de Segmento de Logradouros – CADLOG para os setores urbanos.

O Supervisor, em conjunto com o ACM, participava da escolha e distribuição dos setores correspondentes à área de trabalho sob sua responsabilidade, aos Recenseadores. Iniciada a coleta, sua primeira tarefa era a orientação para que cada Recenseador viesse a conhecer, o melhor possível, o setor onde iria trabalhar. No primeiro dia de coleta de cada setor, o Supervisor preenchia o CD

1.16 – Registro dos Setores Iniciados – de modo que, a cada semana, fossem informados ao ACM os setores distribuídos para coleta.

Era tarefa primordial do Supervisor dar toda a orientação necessária para o trabalho do Recenseador. Assim, durante a primeira semana de coleta do setor, cabia ao Supervisor fazer a distribuição de poucos questionários, buscando evitar que erros de coleta, ocorridos em seu início, pudessem se acumular. Do mesmo modo, nesse período, recomendava-se o comparecimento do Recenseador ao posto de coleta em dois dias distintos, para que os equívocos conceituais e de preenchimento fossem identificados e sanados, evitando-se, assim, que se tornassem sistemáticos.

Ainda na primeira semana de trabalho, o Supervisor acompanhou, em campo, cada um de seus Recenseadores, de modo a avaliar se as instruções para o percurso e cobertura do setor estavam sendo seguidas. Acompanhou, pelo menos, a realização de duas entrevistas, uma onde foi utilizado o Questionário Básico e, outra, o Questionário da Amostra.

A partir da segunda semana de entrevistas, o Supervisor iniciou a rotina de crítica do material coletado, trabalho que se estendeu até o final da coleta do setor. Quando o material era entregue, passava por uma crítica visual na presença do Recenseador. Todos os equívocos identificados, nas folhas de coleta e nos questionários, eram corrigidos no próprio posto ou, se necessário, com o retorno ao campo. O Supervisor controlava, também, a freqüência de domicílios fechados, bem como auxiliava o Recenseador a resolver as situações de recusa.

A rotina de crítica estabelecia ainda que o Supervisor fosse a campo para refazer três entrevistas, avaliasse também a cobertura do setor no trecho que percorresse, bem como verificasse a correta aplicação da fração de amostragem e seleção da amostra. Dependendo do resultado da avaliação desses pontos, a rotina estabelecia que o Recenseador corrigisse os erros encontrados, se necessário com revisão de todo o trabalho da semana. Caso houvesse reincidência em erros graves, o Recenseador passava por retreinamento e, dependendo do seu comportamento, o Supervisor pedia ao ACM a sua substituição.

Era de extrema importância que as informações do CD 1.15 – Agenda do Supervisor – fossem atualizadas. Tudo deveria ser anotado, a produção semanal dos Recenseadores, as observações referentes às rotinas da crítica e o que ainda parecesse importante, para que, de forma organizada, resultasse um histórico dos trabalhos de coleta em cada setor.

Por outro lado, semanalmente, o Supervisor preenchia o CD 1.17 com a produção dos Recenseadores e remetia-o ao ACM, para que fizesse os lançamentos no SIGC. Esse sistema, através do cálculo de parâmetros de controle, fornecia elementos para avaliação do trabalho de coleta, o que determinava verificações, inclusive em campo, tanto para o setor com coleta em andamento ou mesmo após o seu término.

Encerradas as entrevistas no setor, tendo sido criticado o material entregue na última semana e não havendo pendência com o SIGC, o Supervisor conferia o preenchimento do CD 1.06, feito pelo Recenseador, e encerrava a supervisão no setor. Em seguida, empastava os questionários e os remetia ao ACM. O Supervisor aguardava autorização para o empastamento e remessa das Folhas de Coleta do setor.

Sobretudo nos postos não-informatizados, que não dispunham de apoio administrativo próprio, o Supervisor também assumia a responsabilidade de execução das tarefas administrativas relativas à sua equipe. Assim, era responsável pelos procedimentos ligados à contratação de pessoal, incluídos aí a conferência e pedido de alteração de dados cadastrais e o preenchimento de formulários diversos. Cuidava também das adesões ao Programa de Assistência ao Censitário – PAC, bem como das prorrogações e rescisões contratuais.

9.4.3 Trabalho do Agente Censitário Municipal - ACM

O ACM, enquanto responsável pelo posto de coleta, tinha a incumbência de zelar por tudo o que fosse necessário para o bom andamento de todo o trabalho. Isto começava na montagem do posto de coleta, quando verificava as condições físicas do local de trabalho, sobretudo no que se referia à guarda e segurança dos equipamentos e do material de coleta. Além disso, era responsável pela instrução das equipes e pela organização e gerenciamento das atividades de coleta e de revisão da base operacional, bem como pelos procedimentos e controles do Sistema de Administração de Pessoal. Ao ACM também cabia o controle e suprimento de todo o material necessário à rotina do posto de coleta. Nos municípios onde a estrutura censitária determinava apenas um Supervisor, esse cargo era acumulado pelo próprio ACM.

Para o desempenho de todas essas tarefas, o ACM precisava conhecer, além, é claro, das instruções específicas do Manual do Agente Censitário Municipal, todas as instruções contidas nos Manuais do Recenseador, do Supervisor, do Instrutor, para Revisão do Trabalho da Base Operacional Geográfica, do Usuário do Sistema de Indicadores Gerenciais da Coleta e do Sistema de Administração de Pessoal.

Tão logo assumisse suas funções, tomava conhecimento dos setores que compunham a sua área de trabalho, bem como dos quantitativos de Supervisores e Recenseadores que estavam à sua disposição. De posse dessas informações, juntamente com o Coordenador de Subárea, estabelecia as áreas de atuação de cada Supervisor e seus respectivos setores.

Após o treinamento, fazia a chamada dos Recenseadores aprovados e coordenava, em conjunto com os Supervisores, a escolha e distribuição dos setores a serem trabalhados. Durante o período da coleta, também esclarecia os moradores, visando reverter as situações de recusa, que não puderam ser contornadas pelos Recenseadores e Supervisores.

Caso exercesse suas tarefas em um posto informatizado, a operação do SIGC estava sob sua responsabilidade ou a cargo do Supervisor Administrativo, dependendo do que fosse acordado com o Coordenador de Subárea.

Após o encerramento dos trabalhos de campo e liberação do setor pelo SIGC, o ACM encaminhava à agência do IBGE as pastas com os questionários coletados. Caso fosse responsável por um posto não-informatizado, preenchia o CD 1.13.

Após o encerramento de todo o trabalho de coleta do posto, o ACM remetia à agência as Folhas de Coleta utilizadas, fazendo uso, no posto não-informatizado, do CD 1.14. Nessa mesma ocasião, todo o restante do material utilizado – Folha de Domicílio Coletivo, Agenda do Supervisor e Caderneta do Setor – seria também remetido.

9.4.4 Trabalho de coordenação técnica

O Coordenador Técnico e os Coordenadores de Área e Subárea foram os responsáveis diretos pela orientação e gerenciamento do trabalho de campo do censo. O primeiro coordenou todo o trabalho técnico no âmbito da unidade regional, além de fazer a ligação entre a equipe de campo e os órgãos da sede do IBGE. O segundo teve a seu cargo o trabalho de um conjunto de agências e o último comandou os trabalhos na subárea, que, na prática, correspondia à área de atuação da agência do IBGE ou parte dela.

Os coordenadores estabeleciam todo o esquema de execução dos trabalhos de coleta, objetivando a garantia da qualidade das informações e o cumprimento dos prazos estabelecidos. Para tal, orientavam e acompanhavam a execução das rotinas, bem como, através de entendimento com a área responsável na sede, faziam alterações, para melhor ajustá-las à realidade do seu trabalho. Da mesma forma, foram responsáveis pela definição de ações gerenciais específicas, necessárias à correção de qualquer falha que viesse acontecer durante o trabalho de campo.

Dentre as atribuições dos coordenadores, devem ser citadas:

- estabelecer um roteiro de visitas aos postos de coleta;

- distribuir aos postos de coleta todo o material técnico, administrativo e publicitário, necessário à atividade do censo;

- estabelecer uma rotina de supervisão alternativa que permitisse ao Supervisor acompanhar e avaliar a qualidade do trabalho do Recenseador que atuava em setores rurais, de área muito extensa e/ou de difícil acesso;

- definir, juntamente com o ACM, as áreas de atuação dos Supervisores;

- avaliar as correções realizadas pelo ACM nos mapas e nas descrições dos limites dos setores; e

- encaminhar os questionários e as Folhas de Coleta aos Centros de Captura de Dados.

Cabe destacar a responsabilidade dos coordenadores quanto à operação e acompanhamento do SIGC. Inicialmente, teve como atribuição fazer as autorizações de acesso ao sistema, sendo de particular importância a presença semanal dos Coordenadores de Subárea aos postos de coleta não-informatizados, para, com o microcoletor, recolher as informações relativas à produção. Era também de sua alçada atualizar os dados do SIGC nos prazos estabelecidos, bem como consultar diariamente as informações do sistema. Era através desses dados que a equipe de coordenação acompanhava todo o andamento da coleta e, com base nas mensagens geradas, orientava as ações gerenciais necessárias.

No SIGC, os Coordenadores de Subárea estavam autorizados a liberar o pagamento do Recenseador, quando não havia pendência no sistema para o setor trabalhado. No entanto, somente o Coordenador Técnico e o Coordenador de Área podiam autorizar o pagamento de setores pendentes no SIGC, ou dos não-pendentes, que exigiram mais de três retificações em seus dados.

9.5 Acompanhamento da coleta

As informações necessárias ao acompanhamento do trabalho de campo do censo foram produzidas pelo SIGC, pelo Sistema Auxiliar de Indicadores e por uma central de atendimento telefônico. Esses dados, sobretudo das duas primeiras fontes, constituíram a base para a implementação dos mecanismos de acompanhamento e controle em nível local, já descritos no item 9.4.2 e, em nível centralizado, a cargo dos técnicos da Diretoria de Pesquisas – DPE, apresentados mais à frente.

O SIGC permitiu a todos os níveis da estrutura censitária acesso rápido e *on-line* a relatórios sobre o avanço da coleta, oferecendo informações confiáveis e atualizadas, que possibilitavam a cada gestor a tomada de suas decisões.

Já o Sistema Auxiliar de Indicadores mantinha um outro conjunto de informações, montadas a partir do banco de informações do SIGC. Esses dados eram periodicamente atualizados e ficavam disponíveis na página da Coordenação Técnica do Censo Demográfico – CTD, na Intranet da DPE.

No que se refere aos mecanismos de acompanhamento e controle, já a partir da terceira semana de coleta, a DPE colocou em campo um grande número de técnicos que acompanhou os trabalhos em boa parte do País, verificando o preenchimento dos questionários e buscando identificar possíveis falhas conceituais que exigissem urgência na solução.

Numa segunda fase, ocorreram as ações de acompanhamento direcionado da coleta que resultavam de avaliações das informações do SIGC e do Sistema Auxiliar de Indicadores ou, ainda, da solicitação direta das unidades regionais do IBGE. Neste caso, técnicos da sede eram deslocados diretamente para áreas predeterminadas a fim de, em conjunto com as equipes estaduais, darem solução às pendências da coleta.

Durante o trabalho de campo, um sistema de relatórios solicitava aos Coordenadores Técnicos atenção para questões específicas, sobretudo no que tocava ao cumprimento dos prazos, à atualização periódica dos dados no SIGC e à possibilidade de falhas na enumeração da população.

Foi posto em prática, também, um atendimento telefônico que operava sob a responsabilidade do Centro de Documentação e Disseminação de Informações – CDDI, sendo montado a partir do sistema de discagem gratuita 0800, da Embratel. Tal serviço esteve à disposição dos Recenseadores e da população durante o período da coleta, para dirimir dúvidas e registrar reclamações.

9.5.1 Sistema de Indicadores Gerenciais da Coleta

Durante toda a operação de coleta do Censo Demográfico 2000, o IBGE utilizou o Sistema de Indicadores Gerenciais da Coleta – SIGC – como instrumento informatizado de suporte ao controle do trabalho de coleta da pesquisa. agilizando também o processo de pagamento dos Recenseadores, segundo sua produção efetiva.

O SIGC contribuiu para a velocidade na implementação de decisões gerenciais, evitando duplicidade de controles e dados e disponibilizando uma única fonte consolidada de informações para todas as unidades.

9.5.1.1 Descrição do sistema

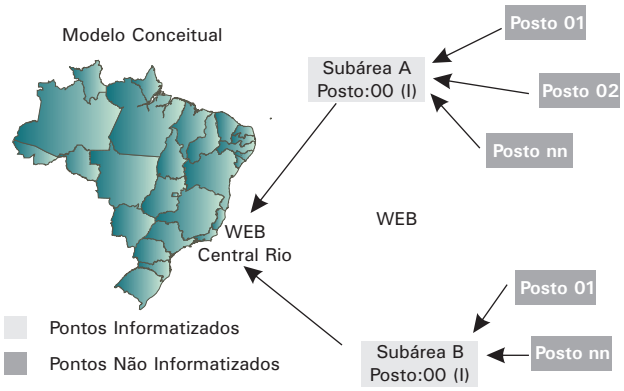
O SIGC foi implementado com tecnologia Lotus Notes / WEB e Crystal Reports em plataformas Windows NT, Unix e Aix, utilizando interface Web e arquitetura Cliente/Servidor.

Para implantação do sistema, foram instalados, aproximadamente, 1000 postos informatizados, cuja localização foi escolhida segundo critérios de facilidade de instalação de recursos de telefonia e condições estruturais para instalação de recursos de informática (segurança, refrigeração, instalação elétrica, etc.).

Cada posto informatizado – ponto fixo – contou com com um microcomputador – arquitetura mínima: Memória de 16 Mb, Disco (HD) de 600 Mb, placa fax modem e softwares Windows 95 e Internet explorer 4.0 – e uma impressora.

Os microcomputadores possuíam acesso ao SIGC, instalado em equipamentos servidores localizados no Rio de Janeiro, através de recursos IP Oficce Dial da Embratel, cuja operação de entrada de dados e recepção de informações eram realizadas pelo ACM, ou outra pessoa autorizada. A figura a seguir, ilustra como era feita a conexão ao sistema.

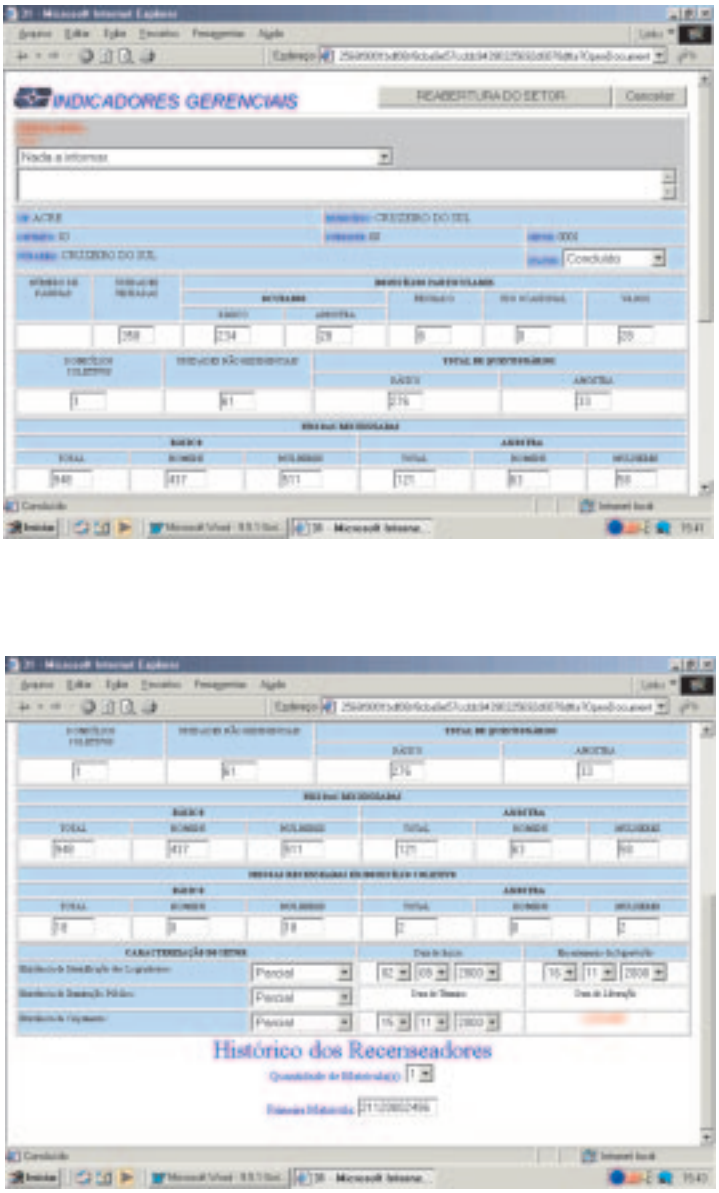
Figura 9.1 – Forma de conexão ao SIGC



Conceitualmente, o SIGC foi desenvolvido e implementado com o objetivo de permitir a entrada de informações para cada um dos setores censitários. Para facilitar o acesso, os setores foram grupados em postos de coleta, e estes em subáreas e unidades regionais. Assim, cada ACM entrava no sistema através de seu posto de coleta infomatizado, dentro da respectiva subárea, após a confirmação de sua senha de acesso, e realizava a atualização das informações relativas à coleta dos setores.

Inicialmente, o ACM informava ao sistema os setores colocados em andamento e as matrículas dos Recenseadores responsáveis pela coleta. Semanalmente, eram atualizadas as informações sobre a coleta do setor – pessoas recenseadas e total de Questionários Básico e da Amostra. Ao final da coleta, o sistema recebia os dados correspondentes ao resumo da Caderneta do Setor. As telas de entrada de dados estão apresentadas a seguir.

Figura 9.2 – SIGC - Telas de entrada de dados para setor com coleta em andamento e concluída

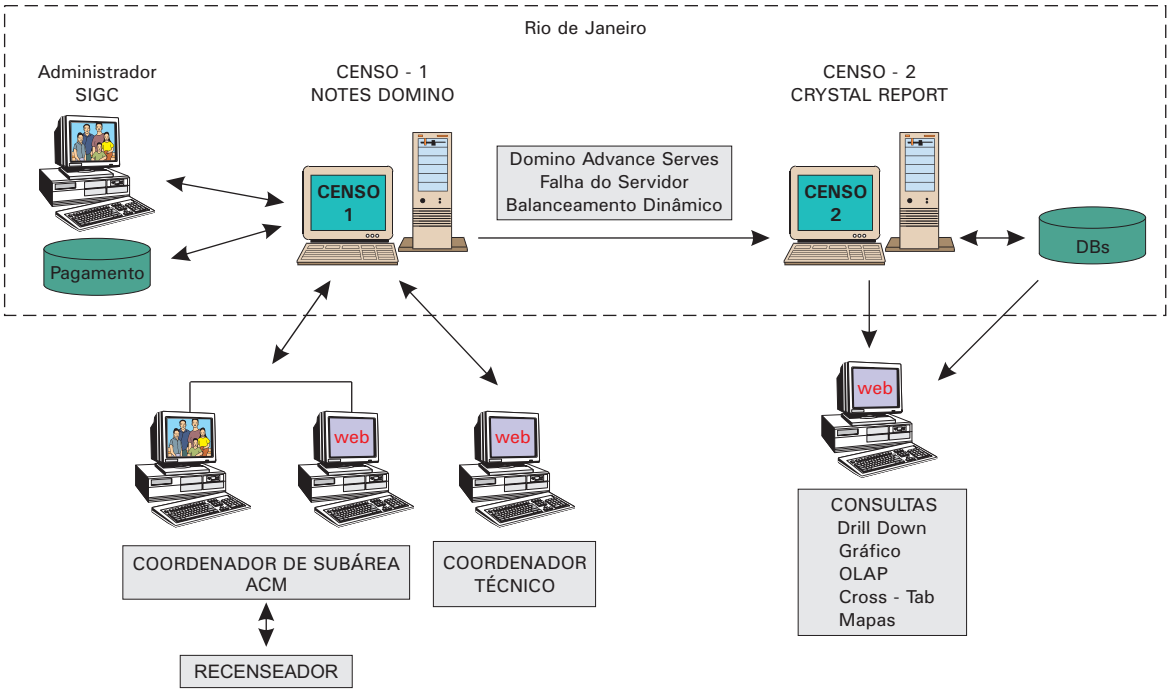


Tela de entrada de dados do SIGC

Fisicamente, foram instalados dois equipamentos servidores no Rio de Janeiro, um para suportar os módulos de atualização Notes dos postos de coleta, e outro para tratar os relatórios gerenciais implementados em Crystal Report informatizados (ver parte superior da figura ilustrativa abaixo).

Foram instalados equipamentos nos postos informatizados com acesso direto ao servidor de atualização Notes e também utilizados equipamentos existentes, incluindo o das Divisões de Pesquisa – DIPEQs, com acesso ao servidor de relatórios gerenciais em Crystal Report (ver parte inferior da figura ilustrativa).

Figura 9.3 – Modelo de instalação do sistema



Foram considerados como pontos móveis do sistema, os postos de coleta não informatizados, ou seja, sem instalação de micro, cujos dados referentes à coleta dos setores foram coletados através de equipamentos móveis, ou seja, coletores portáteis – microcoletores. Esses dados coletados nos pontos móveis foram descarregados nos equipamentos instalados nos postos de coleta – pontos fixos.

O Coordenador de subárea realizava visitas periódicas aos postos não informatizados, segundo agenda previamente comunicada ao ACM, e transcrevia os dados para o microcoletor de dados – cada subárea recebeu um microcoletor. Foram distribuídos, aproximadamente, 1000 microcoletores.

Figura 9.4 - Imagem de um equipamento microcoletor

O SIGC foi implementado num ambiente controlado e estável, com garantia de integridade de dados, identificação dos usuários, autorizações e rastreabilidade de operações, com os seguintes itens de segurança:

- identificação e senhas particulares, definida e mantida pelo usuário;

- bloqueio de acesso por tarefas, estrutura, categoria e funções; e

- gravação automática de registros de controle.

Foi implementada no sistema uma solução que permitia a entrada de dados *off-line* para um conjunto de setores que, em seguida, teriam suas informações atualizadas em bloco, de modo a reduzir o tempo de conexão dos usuários ao sistema.

9.5.1.2 Uso das informações

Sob o ponto de vista do uso gerencial das informações, o SIGC disponibilizou, a partir de sua página principal, dois conjuntos de navegadores: um denominado Consulta em Nível Brasil e outro que congregava as informações para cada uma das 27 unidades da federação.

No navegador Consulta em Nível Brasil, estava à disposição do corpo gerencial do censo um conjunto de relatórios com informações agregadas, segundo o conjunto das unidades da federação, das grandes Regiões e para o Brasil. Era oferecido um retrato ampliado da coleta, através de um conjunto de variáveis, entre as quais merecem destaque:

- número de setores, segundo a situação de coleta, ou seja, não-iniciados, em andamento, concluídos e não-liberados, liberados e paralisados;

- total de pessoas recenseadas, segundo o sexo e por tipo de questionário, para o total de setores, setores em andamento e setores concluídos; e

- número de setores pagos.

No navegador para cada unidade da federação, as variáveis estavam apresentadas segundo níveis hierárquicos de desagregação, começando pela subárea, passando pelo posto de coleta, podendo chegar ao setor. Neste caso, os coordenadores e o ACM utilizaram-se, sobretudo, das seguintes opções: consulta, setores liberados e setores pendentes.

Na opção consulta, foram possíveis duas formas de investigação. A primeira, de título situação de coleta, disponibilizava o quantitativo de setores segundo a situação de coleta, ou seja, nos mesmos moldes apresentados na Consulta em Nível Brasil, com mudança, como foi dito, no nível geográfico de desagregação. A outra forma de investigação, denominada acompanhamento da coleta, fornecia para cada código de setor, segundo o posto ou a subárea, a data de início da coleta, a data da última atualização dos dados, o número de Questionários Básicos e da Amostra utilizados, bem como o número de pessoas recenseadas, por sexo, em cada tipo de questionário.

As outras duas opções, setores liberados e setores pendentes, utilizavam somente a hierarquia subárea, ou seja, apresentavam as respectivas listas de códigos de setores, identificando, assim, o gerente de maior responsabilidade, ligado diretamente ao trabalho de coleta das informações, o Coordenador de Subárea.

a) Indicadores de acompanhamento

Foram definidos dois conjuntos de indicadores para acompanhamento da coleta no setor. No primeiro, três indicadores investigavam os setores em andamento, sendo as avaliações feitas toda vez que houvesse atualização dos dados do setor no SIGC. O outro conjunto, composto de sete indicadores, avaliava o comportamento dos dados do setor após o encerramento da coleta de suas informações, orientando o trabalho final de supervisão.

Foram estabelecidos para cada indicador, separadamente para os setores urbanos e rurais, intervalos de variação para seus valores. Para cálculo dos limites desses intervalos, foram levados em consideração os quartis das distribuições de cada indicador no Censo de 1991 e na Contagem da População de 1996. Inicialmente, essas distribuições foram definidas pelos valores dos setores pertencentes a uma mesma agência de coleta, cuja área, na grande maioria dos casos, coincidia com a da subárea. Exigia-se um mínimo de 50 observações em cada uma das distribuições e, quando isso não se verificava, buscava-se a distribuição para todo o estado ou até para a grande região correspondentes.

Caso o SIGC detectasse divergência entre os dados do setor e os limites estabelecidos, eram emitidas mensagens de advertência – valor do indicador abaixo ou acima do esperado - imediatamente após a atualização dos dados. Essas mensagens exigiam providências imediatas, cabendo ao ACM informá-las aos Supervisores, os quais, após investigação, se necessário, determinavam aos Recenseadores o que fazer.

b) Indicadores para os setores em andamento

No caso dos setores em andamento, foram definidos os seguintes indicadores, somente avaliados pelo sistema quando o setor apresentasse, no mínimo, 50 questionários pesquisados:

- indicador A1 – total de pessoas recenseadas por dia de coleta – objetivava avaliar a produção diária do recenseador, levando em conta o período de coleta do setor; foi calculado como

$$\mathrm{A1}_i = \frac{PA_i + PNA_i}{X_i} \quad i = 1,2,3,\ldots,n$$

onde: n é o número de setores;

PA_i é o número de pessoas recenseadas nos questionários da amostra;

PNA_i é o número de pessoas recenseadas nos questionários da não-amostra; e

X_i é o número de dias de coleta.

Excepcionalmente para esse indicador, os limites de variação – inferior (LI) e superior (LS) – foram fixados a partir de observações do trabalho de campo, nos Censos de 1991 e Experimental, ou seja:

setores urbanos LI = 12, LS = 56 e

setores rurais LI = 9, LS = 42

Caso não existisse um motivo especial, como problemas climáticos, etc., e o Recenseador apresentasse uma produção abaixo do esperado – inferior a LI – cabia ao Supervisor exigir mais esforço tendo em vista melhorar seu desempenho.

Na situação de uma produção acima do valor esperado – superior ao LS – o Supervisor submetia, o mais rápido possível, o material coletado ao plano de supervisão. Se o trabalho apresentasse qualidade, o Recenseador era estimulado a manter o padrão de produção, caso contrário era advertido.

- indicador A2 – diferença entre o número médio de pessoas recenseadas nos Questionários Básicos e nos Questionários da Amostra – investigava diretamente a possibilidade de omissão ou inclusão indevida de pessoas durante as entrevistas; podia também fornecer indícios de equívocos na aplicação da fração de amostragem, ou durante a seleção da amostra; para tanto definiu-se

$$A2'_i = \frac{PNA_i}{K_i} - \frac{PA_i}{L_i} \quad i = 1,2,3,\ldots,n$$

sendo necessário realizar um ajuste, que levasse em conta o número de questionários, resultando em:

$$A2_i = \frac{\sqrt{(K_i \times L_i)}}{\sqrt{(K_i + L_i)}} \times A2'_i$$

onde K_i e L_i são, respectivamente, os números de questionários básicos e da amostra.

Neste caso, os limites de variação foram calculados como

$LI = Q25[A2_{(1991)}] - 2,5 \times (Q75[A2_{(1991)}] - Q25[A2_{(1991)}])$ e

$LS = Q75[A2_{(1991)}] + 2,5 \times (Q75[A2_{(1991)}] - Q25[A2_{(1991)}])$,

onde $Q25[A2_{(1991)}]$ e $Q75[A2_{(1991)}]$ são, respectivamente, o primeiro e o terceiro quartis das distribuições do indicador A2, calculados para os setores concluídos da agência de coleta ou da unidade da federação a qual o setor pertencia, com base nos dados do Censo Demográfico de 1991.

Caso, para um determinado setor, o valor desse indicador estivesse fora dos limites especificados, o Supervisor verificava, de início, se foi feita corretamente a seleção dos domicílios da amostra;

em caso afirmativo, durante a ocasião das reentrevistas, quando da aplicação do plano de supervisão, a possibilidade de omissão ou inclusão indevida de pessoas era investigada com especial atenção; ou

constatados equívocos na seleção dos domicílios da amostra, o Supervisor determinava ao Recenseador que refizesse as entrevistas necessárias, tendo em vista a resolução do problema.

- indicador A3 – razão entre o número de questionários da amostra e o total de questionários; identificava a possibilidade de erro na aplicação da fração amostral, ou de seleção da amostra de domicílios; assim, foi definido

$$A3'_i = \frac{L_i}{(K_i + L_i)} \quad i = 1,2,3,....,n$$

e, a fim de compensar o efeito da maior variabilidade das frações amostrais efetivas, para pequenos setores ou para aqueles com poucos questionários pesquisados, utilizou-se uma transformação, chegando-se a

$$A3_i = \sqrt{(K_i + L_i)} \times \frac{A3'_i - f_i}{\sqrt{f_i \times (1 - f_i)}},$$

onde L_i *e* K_i são, respectivamente, o número de questionários da amostra e da não-amostra, e f_i *é* a fração de amostragem – 10% ou 20% – estabelecida para a coleta do setor; os limites de variação LI e LS foram calculados como

$$LI = Q25[A3_{(1991)}] - 2,5 \times (Q75[A3_{(1991)}] - Q25[A3_{(1991)}]) \text{ e}$$

$$LS = Q75[A3_{(1991)}] + 2,5 \times (Q75[A3_{(1991)}] - Q25[A3_{(1991)}])$$

Ocorrendo para o indicador valores fora dos limites, o Supervisor, inicialmente, investigava se foi feita a correta aplicação da fração de amostragem no setor;

em caso positivo, verificava se ocorreu um número de domicílios fechados, vagos ou de uso ocasional que prejudicou a aplicação do Questionário Básico ou da Amostra; a omissão ou inclusão indevida de pessoas era investigada com especial atenção; ou

constatados erros na aplicação da fração, o Supervisor determinava ao Recenseador que refizesse as entrevistas necessárias, tendo em vista a correção do equívoco.

c) Indicadores para os setores concluídos

Encerrada a coleta do setor e atualizadas suas informações no SIGC, eram calculados os valores para o segundo conjunto de indicadores que, com maior consistência, buscavam investigar, de forma global, o resultado da coleta do setor. Os indicadores foram os seguintes:

- indicador I1 – diferença entre o número médio de pessoas recenseadas nos domicílios particulares da não-amostra e da amostra; a exemplo do que se fez para o indicador A2, definiu-se

$$I1'_i = \frac{PDNA_i}{C_i} - \frac{PDA_i}{D_i} \quad i = 1,2,3,.....,n$$

sendo necessário realizar um ajuste, que levasse em conta o número de domicílios, resultando em:

$$I1_i = \frac{\sqrt{(C_i \times D_i)}}{\sqrt{(C_i + D_i)}} \times I1'_i \text{ onde:}$$

PDAi é o número de pessoas recenseadas nos domicílios particulares da amostra;

$PDNA_i$ é o número de pessoas recenseadas nos domicílios particulares da não-amostra;

D_i e C_i são, respectivamente, os números de domicílios particulares ocupados da amostra e da não-amostra.

Os limites de variação LI e LS foram calculados como

$LI = Q25[I1_{(1991)}] - 2{,}5 \times (Q75[I1_{(1991)}] - Q25[I1_{(1991)}])$ e

$LS = Q75[I1_{(1991)}] + 2{,}5 \times (Q75[I1_{(1991)}] - Q25[I1_{(1991)}])$

- indicador I2 – razão entre o número de domicílios particulares da amostra e o total de domicílios particulares ocupados; a partir de

$$I2'_i = \frac{D_i}{(D_i + C_i)} \quad i = 1,2,3,\ldots,n$$

buscando atenuar o efeito da maior variabilidade das frações amostrais efetivas para pequenos setores, chegou-se a

$$I2_i = \sqrt{(C_i + D_i)} \times \frac{I2'_i - f_i}{\sqrt{f_i \times (1 - f_i)}} ,$$

sendo os limites de variação LI e LS calculados como

$LI = Q25[I2_{(1991)}] - 2{,}5 \times (Q75[I2_{(1991)}] - Q25[I2_{(1991)}])$ e

$LS = Q75[I2_{(1991)}] + 2{,}5 \times (Q75[I2_{(1991)}] - Q25[I2_{(1991)}])$

- indicador I3 – número médio de pessoas recenseadas por domicílio particular ocupado;

$$I3_i = \frac{PDNA_i + PDA_i}{(D_i + C_i)} \quad i = 1,2,3,\ldots,n$$

sendo os intervalos de variação obtidos a partir de

$LI = Q25[I3_{(1996)}] - 2{,}5 \times (Q75[I3_{(1996)}] - Q25[I3_{(1996)}])$ e

$LS = Q75[I3_{(1996)}] + 2{,}5 \times (Q75[I3_{(1996)}] - Q25[I3_{(1996)}])$

- indicador I4 – número de domicílios particulares ocupados; somente foi possível calcular o valor desse indicador, para os setores que se mantiveram inalterados entre 1991 e 2000, ou seja, não foram, nesse período, divididos ou agrupados com outros setores, bem como não tiveram alterações em seus limites.

$I4_i = (D_i + C_i)$

Os intervalos de variação foram definidos a partir do crescimento observado do indicador, entre o Censo de 1991 e a Contagem da População de 1996, ou seja, a partir da razão

$R_{i\ (1996/1991)} = I4_{i(1996)} / I4_{i(1991)}$

corrigida para o período de 4 anos – de 1996 a 2000 – chegando-se a

$R_{i\ (1996/2000)} = (I4_{i(1996)} / I4_{i(1991)})^{4/5}$;

em seguida, a partir das distribuições de $R_{(1996/2000)}$, obteveram-se os limites inferiores LI e superiores LS dessas razões de crescimento tais, como

$LI\ [R_{(1996/2000)}] = Q25[R_{(1996/2000)}] - 2{,}5 \times (Q75[R_{(1996/2000)}] - Q25[R_{(1996/2000)}]$ e

$LS[R_{(1996/2000)}] = Q75[R_{(1996/2000)}] + 2{,}5 \times (Q75[R_{(1996/2000)}] - Q25[R_{(1996/2000)}]$;

assim, chegou-se aos seguintes limites para o indicador I4

$LI = LI\ [R_{(1996/200}] \times I4_{i(1996)}$ e

$LS = LS\ [R_{(1996/200}] \times I4_{i(1996)}$

- indicador I6 – razão entre o número de domicílios fechados e vagos e o total de domicílios particulares ocupados, fechados e vagos

$$I6_i = \frac{F_i + V_i}{(F_i + V_i + D_i + C_i)} \quad i = 1,2,3,....,n$$

onde F_i *e* V_i são, respectivamente, o número de domicílios fechados e vagos; neste caso, tem-se somente um limite superior LS dado por

$LS = Q75[I6_{(1996)}] + 2{,}5 \times (Q75[I6_{(1996)}] - Q25[I6_{(1996)}])$.

- indicador I7 – razão entre o número de domicílios de uso ocasional e o total de domicílios particulares

$$I7_i = \frac{UO_i}{(UO_i + D_i + C_i)} \quad i = 1,2,3,....,n,$$

onde UO_i é o número de domicílios de uso ocasional; como no indicador anterior, temos somente o limite

$LS = Q75[I7_{(1996)}] + 2{,}5 \times (Q75[I7_{(1996)}] - Q25[I7_{(1996)}])$

- indicador I8 – razão de sexo para os domicílios particulares

$$I8_i = \frac{H_i}{(H_i + M_i)} \quad i = 1,2,3,....,n$$

onde H_i e M_i são, respectivamente, o número de homens e de mulheres recenseados nos domicílios particulares; os limites LI e LS foram dados por

$$LI = Q25[I8_{(1996)}] - 2{,}5 \times (Q75[I8_{(1996)}] - Q25[I8_{(1996)}]) \text{ e}$$

$$LS = Q75[I8_{(1996)}] + 2{,}5 \times (Q75[I8_{(1996)}] - Q25[I8_{(1996)}]).$$

Era admitido o status de pendente para um setor com coleta encerrada, no qual um ou mais desses indicadores apresentasse valores fora dos limites estabelecidos; isto, em princípio, impedia que o setor fosse liberado para pagamento. Neste caso, o ACM providenciava a conferência de todos os lançamentos feitos no CD 1.06 – Caderneta do Setor – e da transcrição dessas informações para o SIGC. Caso todas as transcrições e somas estivessem perfeitas, era determinada a volta do setor a campo para a verificação das variáveis cujos comportamentos acarretaram valor(es) fora dos padrões para o(s) indicador(es).

Os gráficos apresentados em seguida mostram a situação de pendência dos setores em relação aos indicadores do SIGC.

Gráfico 9.1 - Setores concluídos pendentes, segundo os indicadores - Brasil

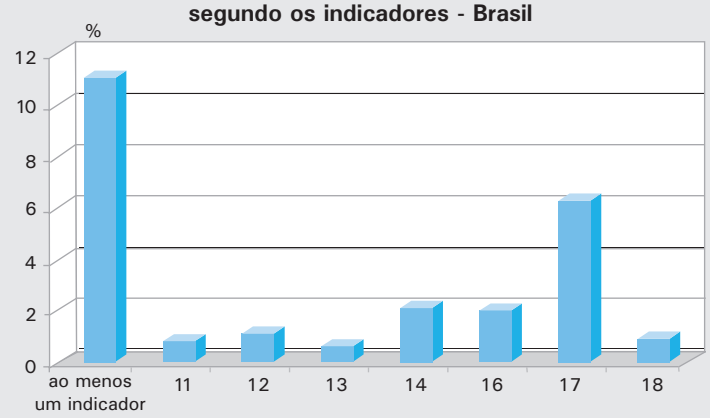
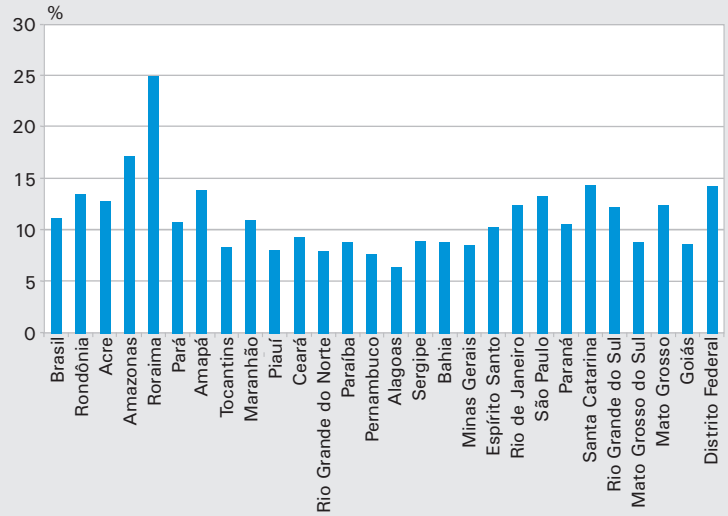


Gráfico 9.2 - Setores concluídos com pendência em pelo menos um indicador, segundo a área de investigação



Na busca de solução das pendências, havendo necessidade de retificação de informações no SIGC, informava-se também se os valores alterados decorreram de investigações com retorno ou não dos técnicos ao campo. Após cada correção, o sistema recalculava os valores dos indicadores.

O pagamento do setor não-pendente com coleta encerrada, ou do que chegou a essa condição após o máximo de três retificações dos seus dados no SIGC, pode ser autorizado pelo Coordenador de Subárea. Já o pagamento de setores pendentes ou liberados, após mais de três alterações, exigia a autorização do Coordenador Técnico ou do Coordenador de Área.

9.5.2 Acompanhamento do trabalho de campo

O trabalho de acompanhamento da coleta do censo foi realizado de duas maneiras: por meio de visitas de técnicos da sede às unidades Regionais e através de um acompanhamento à distância, utilizando relatórios elaborados na DPE, com base nos dados do SIGC e no Sistema Auxiliar de Indicadores, remetidos aos Coordenadores Técnicos nos estados.

9.5.2.1 Sistema Auxiliar de Indicadores

O Sistema Auxiliar de Indicadores foi desenvolvido a partir do banco de informações do SIGC. Para esse sistema, acessado através da página da CTD na Intranet da DPE, montou-se um conjunto de informações periódicas que, além de dar sustentação às ações para gerenciamento do trabalho de campo, contribuía para que, encerrado o censo, fosse possível elaborar uma análise global da coleta das informações.

Esse sistema oferecia os seguintes recursos para o acompanhamento da coleta: informações relativas à Base Operacional Geográfica - BOG; Análise do SIGC; e perguntas e respostas.

O conjunto de informações da BOG tinha como objetivo fornecer suporte ao acompanhamento da coleta. Os valores de suas variáveis eram resultado do trabalho de atualização da base, realizado anteriormente à coleta. Esses dados podiam ser acessados através de um subsistema flexível de consultas, onde se estabeleciam o nível geográfico e um conjunto de variáveis de interesse, cujos valores, se necessário, podiam ser apresentados de forma ordenada. Tinha-se então:

- níveis geográficos – Brasil, grande região, unidade da federação, agência, município, distrito, subdistrito; e

- variáveis – unidades visitadas, domicílios particulares ocupados, número de setores segundo a situação do domicílio e número de setores por tipo especial, não-especial e aglomerado subnormal.

No outro conjunto de informações, denominado Análise do SIGC, era possível, cronologicamente, acessar um conjunto de quadros informativos e também um subsistema de consulta, semelhante ao da BOG, porém mais completo e estruturado. Assim, uma investigação podia ter início através dos quadros informativos e ter continuidade, mais aprimorada, através do auxílio do subsistema de consulta.

A consulta ao conjunto dos quadros informativos – apresentados em anexo –, acontecia com a escolha de uma data, da área de investigação e do quadro desejado. Neste caso, a área de investigação variava, dependendo do

quadro acessado, podendo ser Brasil, grande região, unidade da federação, agência de coleta, município, subárea e posto de coleta.

No que se refere ao subsistema de consulta, era possível, rapidamente, realizar uma pesquisa cronológica para um conjunto específico de variáveis, fazendo-se uso de variáveis de filtro, ou seja, estabelecendo-se critérios de investigação. Por exemplo, para uma unidade da federação e em uma determinada data, podia-se conhecer o número de setores com coleta em andamento e destes os que estavam sem atualização no SIGC há mais de duas semanas, somente nos municípios com mais de 50% dos seus setores com coleta iniciada, ordenados decrescentemente por essa última variável. Também no anexo, tem-se acesso à relação das variáveis incorporadas ao subsistema de consulta.

Por fim, através da opção perguntas e respostas, o Sistema Auxiliar de Indicadores apresentava solução às dúvidas sobre a coleta do Censo. Houve respostas a um total de 95 questões formuladas através do sistema, durante o período de coleta do censo.

9.5.2.2 Visitas às unidades regionais

A presença dos técnicos da sede nas unidades regionais obedeceu ao estabelecido no *Projeto para o acompanhamento da coleta por equipes da sede*: *censo de 2000* (2000) que previa dois modelos ou fases de acompanhamento: o primeiro, de caráter geral, abrangia todas as unidades regionais; o outro, direcionado pelo resultado das análises realizadas na DPE, ou a pedido das próprias unidades, atendia a questões específicas.

As equipes da sede eram constituídas por um conjunto de técnicos, sendo um deles o responsável, cujo efetivo variava de acordo com o volume estimado de trabalho a ser realizado em cada unidade regional.

O modelo de acompanhamento geral foi implementado durante a terceira e quarta semanas de coleta do censo, ainda no mês de agosto de 2000, com a presença de equipes da DPE nas unidades regionais, em média durante duas semanas. Nesse período, em cada unidade, houve acompanhamento obrigatório dos trabalhos no município da capital e em outras áreas do estado, obedecendo a uma proposta de roteiro que podia sofrer alterações em função da sua discussão com os gerentes locais.

Essa fase de acompanhamento tinha como objetivo primordial fazer uma avaliação geral dos trabalhos de coleta, com encaminhamento de soluções. Objetivava, logo no começo dos trabalhos, orientar as equipes de campo, de modo a evitar a sistematização e disseminação de erros, tanto de conceitos como operacionais, ao longo do período de coleta.

O Projeto para o Acompanhamento da Coleta estabelecia um roteiro de investigação que avaliava, em seus pontos mais substantivos, toda a coleta do censo, identificando o que fazer durante a verificação do material no posto de coleta e nas avaliações em campo. Assim, além dos aspectos técnicos ligados à coleta propriamente dita – preenchimento dos instrumentos de coleta, cobertura do setor, seleção da amostra, realização das entrevistas, etc. – deveriam ser avaliadas, no posto de coleta, as condições de trabalho oferecidas, bem como as tarefas ligadas à operação do SIGC. Havia também uma atenção especial à avaliação do trabalho do corpo técnico responsável pelo gerenciamento da coleta, sobretudo dos Supervisores, ACMs e Coordenadores de Subárea, tendo como referência o que foi estabelecido nos respectivos manuais de trabalho.

Ao final dessa primeira fase do acompanhamento, cada técnico elaborava um relatório de suas atividades, sendo, para isso, determinada uma forma padronizada, de modo que o responsável pela equipe de acompanhamento tivesse facilitado o trabalho de consolidação do relatório geral do acompanhamento da coleta na unidade regional.

A segunda fase de acompanhamento da coleta do censo, o acompanhamento direcionado, foi orientada pelos relatórios técnicos da primeira fase e/ou pelo resultado das investigações do Sistema Auxiliar de Indicadores. Podia acontecer também por solicitação das unidades regionais, na grande maioria, para prestar apoio gerencial à coleta, cujos problemas ameaçavam o bom andamento do trabalho, como, por exemplo, a insuficiência do seu quadro de pessoal. Esse acompanhamento aconteceu sempre que necessário e sem prazo definido para execução, o qual variava dependendo do tipo de situação que motivava o acompanhamento direcionado.

Nesse modelo, os técnicos podiam realizar seu trabalho obedecendo a um roteiro prévio determinado na sede ou, no caso de apoio gerencial, a outra modalidade estabelecida em conjunto com a unidade regional. Podiam ser tratadas questões específicas ligadas tanto à coleta em andamento quanto aos trabalhos já concluídos.

Os motivos determinantes do acompanhamento estavam ligados, quase sempre, à produção dos Recenseadores, à aplicação da fração amostral, à pendências em relação aos indicadores do SIGC e aos quantitativos de domicílios visitados no setor e de pessoas recenseadas em nível municipal. No entanto, é bom que se diga que o caráter específico dessa segunda fase do acompanhamento não significava abandonar a observação do trabalho de coleta, em sua totalidade.

9.5.2.3 Sistema de relatórios

Os relatórios tinham como base as informações do SIGC e do Sistema Auxiliar de Indicadores, e eram remetidos pelo Lotus Notes às unidades regionais, sendo mais utilizados a partir do mês de setembro de 2000. Existiam duas modalidades: os relatórios semanais, remetidos para todas as unidades, e os sem periodicidade definida, elaborados em função da identificação de algum problema específico da coleta.

Os relatórios semanais informavam, necessariamente, em nível de unidade da federação e de subárea, a situação dos setores, ou seja, aqueles, com coleta não-iniciada ou paralisada, os sem atualização de dados no SIGC há mais de duas semanas e os com coleta em andamento ou concluída, pendentes por indicadores.

Em nível municipal, o interesse dos relatórios se concentrava no número de domicílios pesquisados e na população recenseada. Eram apresentados aos Coordenadores Técnicos, à medida que a coleta avançava, o número de domicílios particulares ocupados, fechados, vagos e de uso ocasional, bem como o de pessoas recenseadas, além dos resultados correspondentes à Contagem da População de 1996 e às projeções de população para a data de referência do censo. Além dos valores absolutos, eram fornecidas alguns indicadores – quase sempre razões entre variáveis – taxas de crescimento anual e gráficos que davam uma medida da evolução dos números no tempo, facilitando o trabalho gerencial da coleta.

Também em nível de unidade regional, separadamente para as áreas urbana e rural, eram fornecidas, toda semana, as datas previstas para o encerramento da coleta, tendo em vista a produção realizada da unidade até a semana anterior. Complementando esse dado, eram apresentados os números da produção semanal, necessária para que o término do trabalho ocorresse dentro do prazo estabelecido.

9.5.3 Atendimento à sociedade

Durante a coleta das informações, o número 0800-218181, do IBGE, foi reforçado em toda a propaganda e promoção que envolvia o Censo 2000 e também através da divulgação da imprensa por meio de notícias e entrevistas. Na página do Censo 2000 na Internet e em peças promocionais impressas, foi divulgado o *e-mail* censo2000@ibge.gov.br.

O objetivo era levar ao conhecimento da população essas formas de contato com o IBGE, para que as pessoas tivessem à disposição canais de comunicação que lhes permitissem esclarecer dúvidas sobre o censo, confirmar a identificação do Recenseador, reclamar que não haviam sido recenseadas, bem como, dar sugestões, críticas e opiniões.

A estrutura do 0800 contou com até 31 atendentes – nos períodos de pico – distribuídos em quatro turnos, funcionando de 7:00 às 22:00 horas, de domingo a domingo. Havia também uma área responsável pelo acompanhamento e direcionamento dos e-mails às áreas competentes, para respostas aos assuntos específicos de cada mensagem.

Além de servir à população em geral, o serviço de atendimento telefônico 0800 foi utilizado também por funcionários do IBGE e Recenseadores. Dentre seus usos, cabe destacar:

- o registro da entrega de material do censo nas agências, por meio de informações sobre o andamento desse processo;

- o registro dos pedidos de material das agências do IBGE;

- o registro das demandas encaminhadas aos Coordenadores Técnicos do Censo 2000 em cada unidade regional;

- o registro dos domicílios não-recenseados;

- o registro da ocorrência da perda de crachá dos Recenseadores;

- o registro das reclamações dos Supervisores e Recenseadores; e

- a prestação de informações sobre o projeto Vamos Contar!.

De janeiro a dezembro de 2000, o 0800 recebeu aproximadamente 247 mil chamadas. Desse total, mais de 94 000 ligações aconteceram no período da coleta, entre os meses de agosto e dezembro de 2000. Mais de 20% dessas ligações foram referentes a processos seletivos.

Tabela 9.7 - Ligações recebidas mensalmente no 0800 do IBGE - 1999-2000

0800	Ligações recebidas						
	Total	Jan.	Fev.	Mar.	Abr.	Maio	Jun.
1999	11 216	891	655	870	833	1 322	665
2000	247 429	4 622	22 659	21 400	24 769	31 256	20 457

0800	Ligações recebidas					
	Jul.	Ago.	Set.	Out.	Nov.	Dez.
1999	994	1 191	542	738	727	1 788
2000	18 976	37 032	27 253	12 005	15 000	12 000

Fonte: IBGE, Centro de Documentação e Disseminação de Informações, Sistema de Acompanhamento Gerencial do Atendimento.

Com relação ao serviço censo2000@ibge.gov.br, cerca de 400 mensagens foram recebidas por mês no período de agosto a dezembro 2000. Esse *e-mail* registrou solicitações de informações sobre processos seletivos, reclamações da população, dúvidas dos próprios Recenseadores, entre outros itens.

A tabela a seguir mostra as estatísticas de atendimento realizadas pelo 0800 e pelo serviço de e-mail do Censo 2000, detalhadas por tipo de solicitação.

Tabela 9.8 - Estatísticas de atendimento - período 01.08.2000-15.01.2001

Motivo do contato	Total	0800	E-mail
Total	**96 528**	**94 500**	**2 028**
Informação sobre processos seletivos	20 359	20 259	100
Reclamação ou informação solicitada pelo recenseador (pagamento, auxílio-transporte, contra os ACS, entre outros)	11 591	11 333	258
Dúvida do recenseador sobre o preenchimento dos instrumentos de coleta	7 993	7 885	108
Verificação da autenticidade do recenseador	4 695	4 695	-
Reclamação do informante (não foi entrevistado, sobre preenchimento a lápis e assinatura à caneta, falta do quesito deficiente físico, orientação sexual, cor/raça no questionário da amostra, não quer preencher, não quer assinar, entre outros)	4 558	3 301	1 257
Perda ou recuperação de crachá	414	414	-
Reclamação sobre o comportamento do recenseador (sem crachá, mal educado, preenchimento errado, entre outros)	397	369	28
Possibilidade de preencher o questionário pela internet	48	-	48
Outro motivo (trote, engano, ligação não-concluída, Projeto Vamos Contar, agradecimento)	46 473	46 244	229

Fonte: IBGE, Centro de Documentação e Disseminação de Informações, Sistema de Acompanhamento Gerencial do Atendimento.

9.6 Evolução da coleta

Um dos aspectos mais importantes do trabalho de campo, em uma pesquisa do porte de um censo, é o cumprimento do prazo estabelecido para a coleta dos dados. Foi previsto que a maior parte da coleta do Censo 2000 acontecesse no período de 1º de agosto a 31 de outubro, ficando para novembro e dezembro, o trabalho de complemento da coleta.

Este complemento, que consiste na reabertura da coleta em uma parcela dos setores, busca, principalmente, as entrevistas em domicílios fechados, ou seja, naqueles que, apesar de ocupados durante a coleta, não foi possível fazer contato com os moradores. O prazo estabelecido para esse trabalho, respeitava o cronograma do censo, que previa para dezembro de 2000, a divulgação dos Resultados Preliminares.

Assim, antes e durante o período de coleta das informações, foram desenvolvidos esforços no sentido de agilizar o trabalho de campo, com o objetivo de que todos os prazos fossem cumpridos. Pode-se citar, entre outras providências, as ações voltadas para a propaganda da pesquisa junto à população, o cuidado com o treinamento das equipes de campo e a geração de informações que permitissem um acompanhamento gerencial eficiente da evolução da coleta, na sede e nas unidades regionais do IBGE.

Além dessas providências, foram realizadas, no Rio de Janeiro, duas reuniões de trabalho, com a presença de todo o corpo gerencial do censo, da sede, dos Departamentos Regionais e das Divisões de Pesquisa das Unidades Regionais. Tendo como meta específica fazer uma avaliação da coleta, buscou-se, nessas ocasiões, identificar os problemas e encaminhar medidas, visando o cumprimento das metas estabelecidas.

Na primeira reunião, realizada ao final de agosto, que serviu como tomada de posição inicial, foram traçadas para a condução da coleta, as estratégias mais amplas a serem implementadas durante o mês seguinte. Esse tipo de conduta levou em conta que, com menos de 30 dias do início dos trabalhos, ainda não se achavam disponíveis elementos para uma análise e conhecimento mais precisos dos problemas.

A segunda reunião, realizada 30 dias após a primeira, permitiu explicitar a verdadeira situação da coleta para cada unidade regional, sendo possível identificar as dificuldades, técnicas e administrativas, que constituíam os entraves ou "gargalos" ao bom andamento do trabalho. Nessa oportunidade, as providências adotadas, específicas para cada unidade, buscavam fazer com que a conclusão da coleta se afastasse, o menos possível, da prevista no cronograma. Entre as providências adotadas, pode-se citar:

- liberar mais recursos para pagamento de diárias, ajudas de custo e locação de veículos;

- buscar um incremento de auxílio logístico, oferecido por outras Instituições do Governo, como o Exército, a Marinha, prefeituras, etc.;

- contratar novos Recenseadores e remanejar outros, para auxílio aos trabalhos em outros municípios; e

- prorrogar alguns contratos de ACMs e ACSs.

Com todos esses cuidados, a situação do trabalho futuro, sem considerar a atividade de reabertura de setores, foi a seguinte;

- em 31 de outubro, a coleta das informações estava terminada em pouco mais de 80% dos setores; somente a Unidade Regional do Espírito Santo teve a tarefa concluída;

- em 17 de novembro, 12 unidades regionais tinham o trabalho encerrado;

- em 30 de novembro, chegou-se, praticamente, ao término da coleta do censo; e

- em 11 de dezembro, veio a ser efetivamente encerrada a coleta, com o fim dos trabalhos em, aproximadamente, 300 setores restantes, pertencentes a dez unidades da federação.

Uma primeira forma de se observar como evoluiu a coleta do censo é considerar a quantidade de setores, segundo o atributo status da coleta, que classificava o setor nas seguintes categorias: não-iniciado, em andamento, concluído – liberado e não-liberado para pagamento – e paralisado. A tabela abaixo mostra as freqüências para essas categorias.

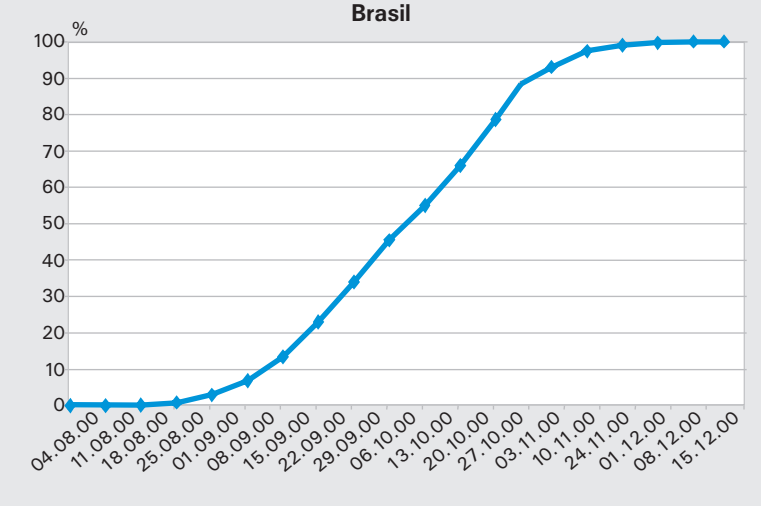
Tabela 9.9 - Total de setores, por status, segundo o período de coleta
Brasil

Período	Não iniciados		Em andamento		Concluídos		Paralisados	
	Absouto	Relativo (%)	Absouto	Relativo (%)	Absouto	Relativo (%)	Absouto	Relativo (%)
04.08.2000	213 782	99,1	1 994	0,9	5	0,0	1	0,0
11.08.2000	180 065	83,4	35 600	16,5	26	0,0	91	0,0
18.08.2000	126 612	58,7	88 318	40,9	298	0,1	554	0,3
25.08.2000	76 703	35,5	135 988	63,0	1 582	0,8	1 509	0,7
01.09.2000	55 269	25,6	151 162	70,1	6 514	3,0	2 837	1,3
08.09.2000	45 365	21,0	152 041	70,5	14 658	6,8	3 718	1,7
15.09.2000	34 634	16,1	148 273	68,7	28 864	13,4	4 011	1,9
22.09.2000	23 146	10,7	139 602	64,7	49 503	23,0	3 531	1,6
29.09.2000	12 785	5,9	126 956	58,8	73 395	34,0	2 646	1,2
06.10.2000	5 149	2,4	111 003	51,4	98 152	45,5	1 478	0,7
13.10.2000	1 809	0,8	94 623	43,9	118 544	55,0	806	0,4
20.10.2000	479	0,2	72 727	33,7	142 262	65,9	314	0,1
27.10.2000	163	0,1	45 914	21,3	169 606	78,6	99	0,0
01.11.2000	50	0,0	25 319	11,7	190 367	88,3	46	0,0
07.11.2000	32	0,0	15 047	7,0	200 677	93,0	26	0,0
14.11.2000	12	0,0	5 290	2,5	210 469	97,5	11	0,0
21.11.2000	1	0,0	2 254	1,0	213 553	99,0	2	0,0
28.11.2000	8	0,0	457	0,2	215 345	99,7	-	0,0
05.12.2000	22	0,0	1	0,0	215 788	100,0	-	0,0
11.12.2000	-	0,0	-	0,0	215 811	100,0	-	0,0

Fonte: IBGE, Censo Demográfico 2000, Sistema de Indicadores Gerenciais da Coleta.

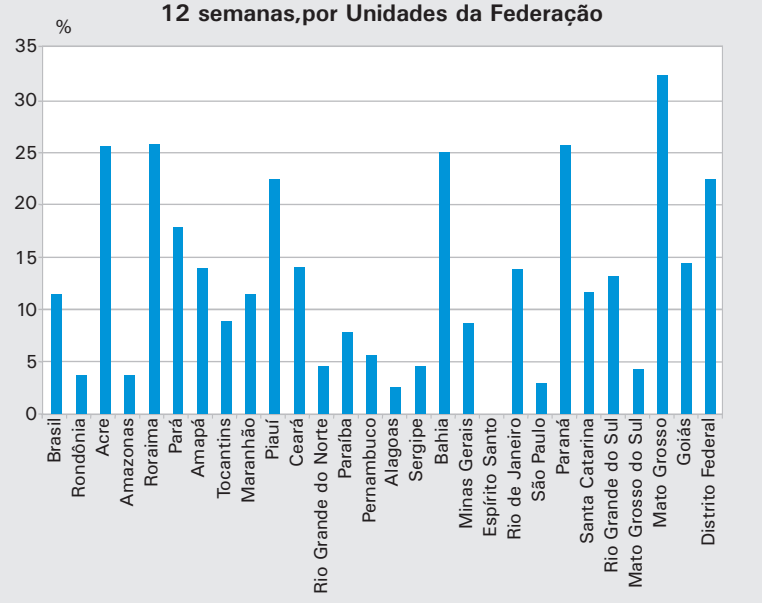
O gráfico 9.3 seguinte ilustra o comportamento do total de setores concluídos ao longo do período.

Gráfico 9.3 - Setores concluídos ao longo do período de coleta Brasil



Uma outra forma de apresentar a evolução da coleta do censo é investigar a freqüência dos setores, segundo o tempo gasto para a conclusão da coleta. Assim, por exemplo, a situação dos setores, cujos tempos de coleta superaram 12 semanas, é apresentada no gráfico seguinte, segundo as Unidades da Federação.

Gráfico 9.4 - Setores, cujos tempos na coleta superaram 12 semanas,por Unidades da Federação



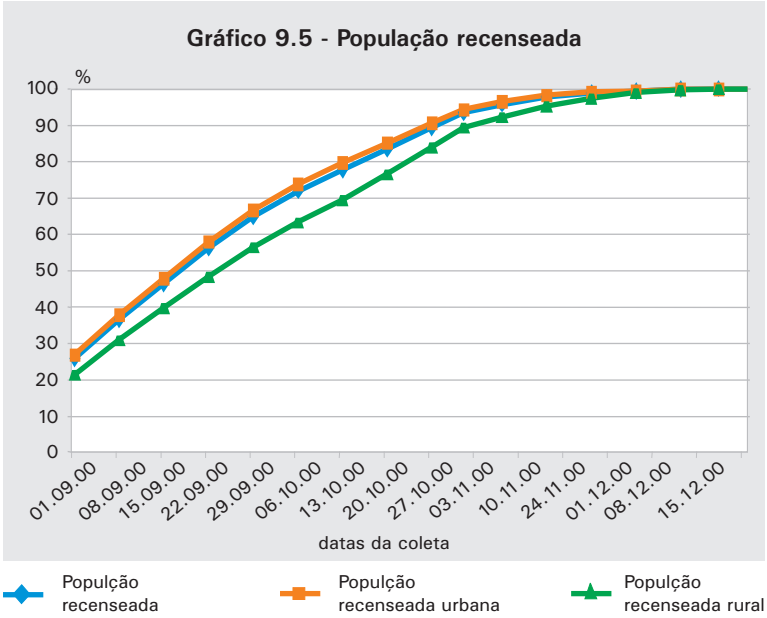
A tabela a seguir mostra o resultado completo da investigação da duração da coleta dos setores, para o Brasil e as unidades da federação.

Tabela 9.10 - Total de setores, por tempo gasto em semanas para a conclusão da coleta do setor, segundo as Unidades da Federação

Unidades da Federação	Tempo gasto em semanas para a conclusão da coleta do setor										
	Total	Até 2		Mais de 2 a 4		Mais de 4 a 8		Mais de 8 a 12		Mais de 12	
		Absoluto	Relativo (%)	Absoluto	Relativo (%)	Absoluto	Relativo (%)	Absoluto	Relativo (%)	Absoluto	Relativo (%)
Brasil	**215 811**	**17 878**	**8,28**	**38 458**	**17,82**	**83 559**	**38,72**	**51 429**	**23,83**	**24 487**	**11,35**
Rondônia	1 996	457	22,90	503	25,20	631	31,60	333	16,70	72	3,60
Acre	555	39	7,00	75	13,50	186	33,50	114	20,60	141	25,40
Amazonas	3 317	374	11,30	803	24,20	1 433	43,20	589	17,80	118	3,60
Roraima	507	58	11,44	63	12,43	173	34,12	83	16,37	130	25,64
Pará	6 148	437	7,10	877	14,30	2 157	35,00	1 591	25,90	1 086	17,70
Amapá	463	67	14,50	48	10,40	160	34,50	124	26,80	64	13,80
Tocantins	1 367	117	8,60	200	14,60	477	34,90	453	33,10	120	8,80
Maranhão	6 447	633	9,80	1 022	15,90	2 004	31,10	2 057	31,90	731	11,30
Piauí	3 717	197	5,30	539	14,50	1 142	30,70	1 011	27,20	828	22,30
Ceará	7 965	595	7,50	1 376	17,30	2 972	37,30	1 917	24,10	1 105	13,90
Rio Grande do Norte	2 655	138	5,20	263	9,90	882	33,20	1 252	47,20	120	4,50
Paraíba	4 171	244	5,80	596	14,30	1 645	39,50	1 365	32,70	321	7,70
Pernambuco	8 558	345	4,00	1 079	12,60	3 378	39,50	3 288	38,40	468	5,50
Alagoas	2 610	127	4,90	234	9,00	855	32,70	1 330	51,00	64	2,50
Sergipe	2 222	87	3,92	303	13,64	1 030	46,35	703	31,64	99	4,46
Bahia	15 342	552	3,60	1 585	10,33	4 731	30,84	4 662	30,39	3 812	24,85
Minas Gerais	22 544	1 639	7,30	3 830	17,00	9 305	41,20	5 827	25,90	1 943	8,60
Espírito Santo	3 228	226	7,00	488	15,10	1 485	46,00	1 028	31,90	1	0,00
Rio de Janeiro	20 756	1 491	7,20	4 462	21,50	8 736	42,10	3 222	15,50	2 845	13,70
São Paulo	49 713	4 520	9,10	11 127	22,40	23 168	46,60	9 457	19,00	1 441	2,90
Paraná	13 060	1 134	8,70	1 769	13,50	3 703	28,40	3 122	23,90	3 332	25,50
Santa Catarina	6 833	573	8,40	1 212	17,70	2 498	36,60	1 767	25,90	783	11,50
Rio Grande do Sul	16 907	2 236	13,20	3 635	21,50	5 888	34,80	2 933	17,40	2 215	13,10
Mato Grosso do Sul	2 714	396	14,60	596	22,00	1 098	40,50	511	18,80	113	4,20
Mato Grosso	3 377	536	15,90	498	14,70	728	21,60	528	15,60	1 087	32,20
Goiás	5 966	422	7,10	927	15,50	2 128	35,70	1 636	27,40	853	14,30
Distrito Federal	2 673	238	8,90	348	13,02	966	36,14	526	19,68	595	22,26

Fonte: IBGE, Censo Demográfico 2000, Sistema de Indicadores Gerenciais da Coleta.

Por fim, é interessante também observar a evolução da coleta, tendo em conta o percentual da população recenseada. O gráfico seguinte mostra essa informação, pesquisada a partir de 1º de setembro, segundo a situação do domicílio.



O trabalho de reabertura da coleta de um setor era realizado pelo Supervisor, sendo para isso, se necessário, renovado seu contrato de prestação de serviços.

Essa atividade teve início tão logo o setor fosse liberado para pagamento pelo SIGC, e compreendeu, sobretudo, a obtenção de entrevistas nos domicílios classificados como fechados durante o período de coleta, mas também, para a correção de erros, como por exemplo, de omissão na coleta de domicílios, ou sua inclusão indevida no setor, resultado de invasão de outras áreas de coleta.

O impacto do trabalho de reabertura de setores é apresentado a seguir, em cada unidade da federação, levando em conta o número de setores, de domicílios e de pessoas recenseadas.

Tabela 9.11 - Impactos da reabertura da coleta de setores no censo para algumas variáveis, segundo as Unidades da Federação

(continua)

Unidades da Federação	Setores			Domicílios fechados			Pessoas recenseadas		
	Total	Reabertos	Relativo (%)	Total	Reabertos	Relativo (%)	Total	Reabertos	Relativo (%)
Brasil (1)	**215 811**	**44 549**	**20,64**	**754 113**	**225 439**	**29,89**	**169 303 227**	**848 351**	**0,50**
Rondônia	1 996	300	15,03	7 398	1 174	15,87	1 377 545	3 731	0,27
Acre	555	127	22,88	7 473	803,00	10,75	555 011	3 309	0,60
Amazonas	3 317	350	10,55	15 296	1 912	12,50	2 839 261	(-) 9 787	(-) 0,34
Roraima	507	139	27,42	3 091	691,00	22,36	321 960	5 221	1,62
Pará	6 148	479	7,79	39 819	2 842	7,14	6 191 810	8 159	0,13
Amapá	463	49	10,58	5 112	437,00	8,55	474 236	3 207	0,68
Tocantins	1 367	261	19,09	4 102	929,00	22,65	1 154 218	4 975	0,43
Maranhão	6 447	465	7,21	25 486	2 916	11,44	5 638 271	11 495	0,20
Piauí	3 717	316	8,50	4 976	937,00	18,83	2 840 507	4 068	0,14
Ceará	7 965	1 052	13,21	22 052	3 823	17,34	7 406 528	22 637	0,31
Rio Grande do Norte	2 655	97	3,65	14 870	1 366	9,19	2 775 268	523	0,02
Paraíba	4 171	572	13,71	6 632	1 449	21,85	3 435 621	9 364	0,27
Pernambuco	8 558	1 572	18,37	24 923	6 394	25,66	7 900 513	29 075	0,37
Alagoas	2 610	394	15,10	10 097	2 592	25,67	2 814 319	8 882	0,32

Tabela 9.11 - Impactos da reabertura da coleta de setores no censo para algumas variáveis, segundo as Unidades da Federação

(conclusão)

Unidades da Federação	Setores			Domicílios fechados			Pessoas recenseadas		
	Total	Reabertos	Relativo (%)	Total	Reabertos	Relativo (%)	Total	Reabertos	Relativo (%)
Sergipe	2 222	358	16,11	9 251	1 740	18,81	1 779 221	5 248	0,29
Bahia	15 342	4 879	31,80	56 889	24 837	43,66	12 945 611	147 677	1,14
Minas Gerais	22 544	3 259	14,46	46 921	13 404	28,57	17 873 446	50 858	0,28
Espírito Santo	3 228	1 142	35,38	23 524	14 511	61,69	3 056 947	44 089	1,44
Rio de Janeiro	20 756	6 967	33,57	109 814	42 635	38,82	14 291 614	127 497	0,89
São Paulo	49 713	8 581	17,26	196 689	45 459	23,11	36 989 306	158 696	0,43
Paraná	13 060	3 652	27,96	22 874	11 624	50,82	9 547 926	43 759	0,46
Santa Catarina	6 833	2 206	32,28	16 918	11 198	66,19	5 319 023	46 317	0,87
Rio Grande do Sul	16 907	3 660	21,65	14 808	8 232	55,59	10 192 620	32 157	0,32
Mato Grosso do Sul	2 714	904	33,31	7 627	3 990	52,31	2 071 108	12 909	0,62
Mato Grosso	3 377	1 212	35,89	15 592	7 978	51,17	2 487 740	37 010	1,49
Goiás	5 966	939	15,74	25 249	8 966	35,51	4 981 842	28 994	0,58
Distrito Federal	2 673	617	23,08	16 630	2 600	15,63	2 041 755	8 281	0,41

Fonte: Censo 2000 – Sistema de Indicadores Gerenciais da Coleta.
(!) Totais lançados no SIGC ao término da coleta, antes de iniciado o trabalho de reabertura dos setores.

Referências

ALBIERI, S. *Uma discussão sobre a fração amostral na coleta dos dados dos censos demográficos brasileiros.* Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1998. 12 p.

_____; BIANCHINI, Z. M. *Sobre o uso de amostragem de setores na coleta dos dados dos censos demográficos brasileiros.* Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1998. 12 p.

BIANCHINI, Z. M.; ALBIERI, S. *Amostragem na coleta dos dados do censo demográfico 2000:* uma versão resumida. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1999a. 9 p.

_____. *Estudo de alternativas para as frações de amostragem do censo demográfico 2000.* Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1999b. 8 p.

BRANDÃO, A. L. S.; LIMA, J. M.; COSTA, L. N.; SILVA, P. L. N. *Sobre a influência dos quesitos pesquisados no esquema de amostragem do censo demográfico:* versão preliminar. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1987. 29 p.

BRASIL. Lei no 8.443, de 16 de julho de 1992. Dispõe sobre a lei orgânica do Tribunal de Contas da União e dá outras providências. *Diário Oficial [da] República Federativa do Brasil,* Poder Executivo, Brasília, DF, 16 jul. 1992. p. 9449. Col. 1.

CENSO demográfico 2000: estudo sobre postos de coleta: (texto para discussão). Rio de Janeiro: IBGE, 1999. 12 p.

_____: instruções para revisão do trabalho da base operacional geográfica – TR 23. Rio de Janeiro: IBGE, 2000. 32 p.

CENSO demográfico 2000: manual de coordenação técnica – CD 1.12. Rio de Janeiro: IBGE, 2000. 19 p.

_____: manual do agente censitário municipal – CD 1.11. Rio de Janeiro: IBGE, 2000. 33 p.

_____: manual do instrutor – TR 01. Rio de Janeiro: IBGE, 2000. 244 p.

_____: manual do instrutor – TR 18. Rio de Janeiro: IBGE, 2000. 72 p.

_____: manual do recenseador – CD 1.09. Rio de Janeiro: IBGE, 2000. 151 p.

_____: manual do sistema de administração de pessoal: orientações gerais – AD 01. Rio de Janeiro: IBGE, 2000. 45 p.

_____: manual do supervisor – CD 1.10. Rio de Janeiro: IBGE, 2000. 25 p.

_____: manual do usuário do sistema de indicadores gerenciais da coleta de dados – TR 24. Rio de Janeiro: IBGE, 2000. 37 p.

COCHRAN, W. G. *Sampling techniques.* 3rd ed. New York: Wiley, c1977. 428 p.

COSTA, L. N. Aplicação da amostragem na coleta dos censos demográficos no Brasil. *Revista Brasileira de Estatística,* Rio de Janeiro, v. 48, n. 189/190, p. 35-64, 1987.

LIMA, J. M.; COSTA, L. N. *Avaliação da consistência das estimativas de características de domicílios obtidas através de uma amostra de 10% simulada a partir da amostra de 25% do CD-80.* Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1988. 18 p.

METODOLOGIA do censo demográfico de 1980. Rio de Janeiro: IBGE, 1983. 478 p. (Série relatórios metodológicos , v. 4).

PINHEIRO, J. C. da R. C.; LIMA, J. M. de L. *Avaliação dos efeitos de redução da fração de amostragem no censo demográfico.* Rio de Janeiro: IBGE, Escola Nacional de Ciências Estatísticas, 1988. 48 p. (Relatórios técnicos, 4/88).

_____. _____. *Revista Brasileira de Estatística,* Rio de Janeiro, v. 51, n. 195, p. 53-84, 1990.

_____. *Uma proposta de dimensionamento da amostra do censo demográfico de 1990.* Rio de Janeiro: IBGE, Escola Nacional de Ciências Estatísticas, 1989. 49 p. (Relatórios técnicos, 4/89).

PROJETO de avaliação das atividades de planejamento e da coleta do censo 2000. Rio de Janeiro: IBGE, Coordenação Operacional dos Censos, 2002. v. 4, 6 e relatório final.

PROJETO para o acompanhamento da coleta por equipes da sede: censo de 2000. Rio de Janeiro: IBGE, 2000. 12 p.

SILVA, L. F.; BIANCHINI, Z. M. *A redução da amostra e a utilização de duas frações amostrais no censo demográfico de 1980.* Rio de Janeiro: IBGE, 1990. 49 p. (Textos para discussão, n. 33).

SILVA, P. L. do N.; MOURA, F. A. da S. Efeito de conglomeração da malha setorial do censo demográfico de 1980. Rio de Janeiro: IBGE, 1990. 115 p. (Textos para discussão, n. 32).

_____; PESSOA, D. G. C.; CARDOSO, R. L.; AUGUSTO, M. de F. L. Censo demográfico 2000: detecção de setores com dados suspeitos no controle da coleta. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1999. 11 p.

_____; PESSOA, D. G. C.; SANTOS, A. dos R. Detecção de valores extremos para fração amostral observada. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2000. 3 p.

SKINNER, C. J.; HOLT, D.; SMITH, T. M. F. (Ed.) Analysis of complex surveys. New York: Wiley, c1989. 309 p.

Anexos

Anexo 9.1 Descrição dos Quadros do Sistema Auxiliar de Indicadores

Quadro 1 – Setores por status na coleta, segundo a área de investigação e a situação do domicílio.

Quadro 3 – Setores em andamento por período sem atualização, segundo a área de investigação e a situação do domicílio.

Quadro 4 – Setores em andamento por tempo de duração da coleta, segundo a área de investigação e a situação do domicílio.

Quadro 5 – Setores e setores concluídos ou liberados por tempo de duração da coleta, segundo a área de investigação e a situação do domicílio.

Quadro 6 – Setores liberados por tempo de liberação (entre a data de liberação e a data de encerramento da supervisão), segundo a área de investigação e a situação do domicílio.

Quadro 6A – Setores liberados por tempo de liberação (entre a data de liberação e a data de término da coleta), segundo a área de investigação e a situação do domicílio.

Quadro 7 – Setores liberados por condição de liberação, segundo a área de investigação e a situação do domicílio.

Quadro 8 – Setores em andamento com total de questionários investigados igual ou superior a 50, pendentes por indicadores, segundo a área de investigação e a situação do domicílio.

Quadro 8A – Setores em andamento com atualizações no período, com total de questionários investigados igual ou superior a 50, pendentes por indicadores, segundo a área de investigação e a situação do domicílio.

Quadro 9 – Setores concluídos ou liberados por tipo de ocorrência e pendência por indicador, segundo a área de investigação e a situação do domicílio.

Quadro 10 – Percentual de setores concluídos ou liberados por tipo de ocorrência e pendência por indicador, segundo a área de investigação e a situação do domicílio.

Quadro 12 – Setores em andamento com atualizações no período, estatísticas básicas da produtividade média diária (pessoas entrevistadas por dia) não-acumulada, segundo a área de investigação e a situação do domicílio.

Quadro 12A – Setores em andamento com atualizações no período, média da produtividade média diária (pessoas entrevistadas por dia) não-acumulada, segundo a área de investigação e a situação do domicílio.

Quadro 13 – Setores em andamento com atualizações no período, estatísticas básicas da produtividade média diária (pessoas entrevistadas por dia) acumulada, segundo a área de investigação e a situação do domicílio.

Quadro 15 – População projetada, população recenseada, setores e domicílios particulares, segundo a área de investigação.

Quadro 16 – População projetada, população recenseada, setores concluídos ou liberados, setores e domicílios particulares, segundo a área de investigação.

Quadro 19A – Projeções para encerramento da coleta por produtividade média diária (pessoas entrevistadas por dia), segundo a área de investigação.

Quadro 20 – Taxa anual de crescimento, variação da população e razão de sexo para municípios concluídos, segundo a área de investigação e a situação do domicílio.

Quadro 21 – Taxa anual de crescimento e variação de domicílios para municípios concluídos, segundo a área de investigação e a situação do domicílio.

Quadro 22 – Taxa anual de crescimento, variação da população e razão de sexo para municípios liberados, segundo a área de investigação e a situação do domicílio.

Quadro 23 – Taxa anual de crescimento e variação de domicílios para municípios liberados, segundo a área de investigação e a situação do domicílio.

Anexo 9.2 Variáveis disponíveis na consulta ao Sistema Auxiliar de Indicadores

Total de Setores

Setores não-iniciados

Setores em andamento

Setores paralisados

Setores concluídos

Setores liberados

Proporção de Setores não-iniciados

Proporção de Setores em andamento

Proporção de Setores paralisados

Proporção de Setores concluídos

Proporção de Setores liberados

Set em andamento com mais de duas semanas sem atualização

% de set em andamento com mais de duas semanas sem atualização

Set em andamento com mais de cinco semanas de coleta

% de set em andamento com mais de cinco semanas de coleta

Setores liberados com tempo de liberação superior a duas semanas

Setores liberados por comando

% de Setores liberados por comando

Setores em And com Pendência em pelo menos um Indicador

% de Setores em And com Pendência em pelo menos um Indicador

Setores Concluídos com Pendência em Pelo Menos um Indicador

% de Setores Concluídos com Pendência em pelo menos um Indicador

Total de Setores Concluídos ou Liberados com Codfor Menor que 200

Setores Concluídos Codfor Menor que 200 com Pendência pelo I4

Média da Prod não Acum Set em And com Atualiz no Período

CV da Prod não Acum Set em And com Atualiz no Período

Média da Prod Acumulada Setores em And com Atualiz no Período

CV da Prod Acumulada Set em And com Atualiz no Período

Setores Concluídos com Diferença de Prod Superior a 20%

% de Setores Conc com Diferença de Prod Superior a 20%

Dif Relativa entre a Pop Projetada e a Pop Recenseada

Dif Relativa entre o Total de Setores e Setores Conc + And

Dif Relativa entre BOG e DPOpesq + Quest Bás e Amostra

Média da Diferença Relativa entre BOG e DPOpesquisados

CV da Diferença Relativa entre BOG e DPOpesquisados

Média da Diferença Relativa entre BOG e DPOpesquisados

Média da Diferença Relativa entre Pop Prevista e Pesquisada

CV da Diferença Relativa entre Pop Prevista e Pesquisada

Média da Diferença Relativa entre Pop Prevista e Pesquisada

Total de Dias Previstos até o Encerram da Coleta Segundo DPO

Total de Dias Previstos até o Encerram da Coleta Segundo POP

Set em andamento com até uma semana sem atualização

Set em andamento com + de uma a duas semanas sem atualização

% de set em andamento com até uma semana sem atualização

% de set em andamento com + de uma a duas semanas sem atualização

Setores em Andamento com mais de 50 questionários

% de Setores em Andamento com mais de 50 questionários

Anexo

CD 1.01 Questionário Básico do Censo Demográfico 2000

CD 1.02 Questionário da Amostra do Censo Demográfico 2000

CD 1.03 Folha de Domicílio Coletivo do Censo Demográfico 2000

CD 1.06 Caderneta do Setor do Censo Demográfico 2000

CD 1.07 Folha de Coleta do Censo Demográfico 2000

10 Pesquisa de Avaliação da cobertura da coleta

É natural que na execução de uma pesquisa como o censo demográfico ocorram algumas falhas de cobertura, sobre tudo quando realizada em um país como o Brasil que apresenta, em seu vasto território, situações bastante difíceis para a coleta de informações. Assim, é importante que se conheça, tanto quanto possível, a magnitude dessas falhas a fim de subsidiar a interpretação dos resultados e o aperfeiçoamento de futuros censos.

A mensuração do erro de cobertura pode ser feita através de análises demográficas, estudos comparativos com informações externas ao censo e, também, através da realização de uma pesquisa específica e independente do censo, que avalie a cobertura de sua coleta.

Desde o Censo Demográfico 1970, o IBGE vem se preocupando com a avaliação da coleta, especialmente com o intuito de conhecer o grau de cobertura alcançado, tendo, para isso, elaborado indicadores que permitem avaliar a enumeração de domicílios bem como de pessoas, conforme visto em Costa (1990).

Algumas fontes que levam à ocorrência de falhas de cobertura são a interpretação incorreta de conceitos, definições ou procedimentos estabelecidos nos manuais técnicos, problemas na definição dos limites dos setores censitários etc. Algumas das possíveis conseqüências dessas falhas são:

- omissão de unidades residenciais ou não-residenciais; e

- omissão de pessoas residentes em domicílios particulares ocupados.

É importante ser ressaltado que essas falhas não são passíveis de correção após encerrada a coleta do censo; é nesse contexto que se insere a Pesquisa de Avaliação da Cobertura da Coleta do Censo Demográfico 2000 – PA. Realizada por amostragem, compreendeu uma segunda coleta em setores selecionados, em cada uma das Unidades da Federação. Ocorreu após o final da coleta de cada setor pelo censo, que teve suas informações confrontadas com as coletadas pela pesquisa.

10.1 Planejamento da pesquisa

A PA foi testada anteriormente, com a realização da Pesquisa de Avaliação do Censo Experimental 1999, durante a qual foi possível verificar todo o trabalho da pesquisa, desde o treinamento das equipes, passando pela coleta das informações, até as fases do Batimento – confronto das informações – e da Reconciliação das Informações – retorno aos domicílios para esclarecimentos adicionais, conforme descrito no *Relatório da pesquisa de avaliação do censo experimental* (1999).

O método escolhido para estimação dos indicadores do grau de cobertura do censo foi o *Dual System Estimation*, cujas premissas básicas determinam que a coleta da PA seja realizada nos mesmos moldes que a do censo e que haja total independência entre as duas pesquisas. Essa independência deve acontecer tanto em relação ao material da coleta do censo, que não pode ser consultado durante a coleta da PA, quanto às equipes de coleta das duas pesquisas.

Com o intuito de garantir essa independência, decidiu-se utilizar na PA uma equipe exclusiva de Recenseadores e Supervisores, recrutada com base na lista do pessoal aprovado no teste seletivo realizado para a coleta do Censo Demográfico 2000, mas não aproveitado durante essa operação. Optou-se também por dar início à coleta da PA ainda durante a coleta do censo, buscando reduzir a chance de perda de informação por motivo de mudança de famílias que ocupassem domicílios no Censo e neles já não se encontrassem por ocasião da PA.

Foi criada a função de Coordenador Estadual da Pesquisa de Avaliação, ocupada por um técnico do quadro efetivo do IBGE que, desligado dos trabalhos do censo a partir de sua nomeação, teve a função de gerenciar, administrativa e tecnicamente, todo o trabalho de coleta, do batimento e da reconciliação.

Para o treinamento das equipes da pesquisa, estabeleceu-se como critério, primeiramente, fazer a capacitação para a coleta dos dados e, mais tarde, dar a instrução para as fases de Batimento e Reconciliação das Informações.

A PA foi realizada em todo o território nacional, com exceção dos setores rurais pertencentes à Região Norte, por insuficiência de recursos. Foram avaliadas a cobertura dos domicílios particulares ocupados e das pessoas neles residentes, situados nos setores não-especiais urbanos e rurais. Considerando-se a possibilidade de serem selecionados setores com número de domicílios muito reduzido, o que provavelmente viria a prejudicar os objetivos da PA, foi estabelecido que apresentassem, no mínimo, 20 unidades visitadas no censo.

Desenvolveu-se, também, o Sistema de Acompanhamento da Coleta, Batimento e Reconciliação, que deu o necessário suporte de informática à pesquisa, desde a coleta das informações até o cálculo dos indicadores de cobertura do censo.

10.2 Amostra da pesquisa

Oliveira, Freitas e Bianchini (1999) elaboraram o documento *Definição do desenho amostral*, baseado em um estudo desenvolvido pela DPE com o intuito de se definir o desenho amostral a ser adotado para a PA. De acordo com esse estudo, foi recomendada a utilização do esquema de amostragem de conglomerados em dois estágios de seleção – setores e domicílios – em cada Unidade da Federação; ou seja, considerou-se o setor censitário como sendo um conglomerado de domicílios.

O plano da amostra da PA foi elaborado com a finalidade de fornecer, com base nas unidades de primeiro estágio – setores – estimativas para o Brasil e Unidades da Federação, para os seguintes indicadores:

- taxas de omissão do número de unidades visitadas, do número total de domicílios e do número de domicílios particulares;

- taxas de omissão do número de pessoas residentes em domicílios particulares ocupados;

- taxas de inclusão indevida de domicílios particulares ocupados; e

- taxas de inclusão indevida de pessoas.

Os setores foram selecionados com eqüiprobabilidade, de forma sistemática, em cada Unidade da Federação, após a realização de uma reordenação nos arquivos da base operacional. Isso objetivou separar os setores em urbanos e rurais, existentes em cada distrito ou subdistrito de cada município pertencente a cada microrregião, de modo a garantir sua representatividade na amostra da pesquisa.

Cabe ressaltar que o dimensionamento da amostra de setores foi elaborado com vistas ao controle dos erros de amostragem, referentes às estimativas das taxas de omissão em nível do total, já que se fosse levada em conta a situação do domicílio, urbano ou rural, provavelmente seria necessário um aumento considerável no tamanho da amostra, que permitisse gerar as estimativas correspondentes com níveis de precisão adequados. A Tabela 10.1 a seguir informa sobre o número de unidades de primeiro estágio da amostra da PA.

Tabela 10.1 - Número de setores censitários, selecionados para a Pesquisa de Avaliação, por situação do domicílio, segundo as Unidades da Federação - 2000

(continua)

Unidades da Federação	Número de setores censitários, por situação do domicílio		
	Total	Urbano	Rural
Brasil	**1354**	**994**	**360**
Rondônia (1)	28	28	...
Acre (1)	26	26	...
Amazonas (1)	40	40	...
Roraima (1)	26	26	...
Pará (1)	40	40	...
Amapá (1)	26	26	...
Tocantins (1)	40	40	...
Maranhão	44	18	26

Tabela 10.1 - Número de setores censitários, selecionados para a Pesquisa de Avaliação, por situação do domicílio, segundo as Unidades da Federação - 2000

(conclusão)

Unidades da Federação	Número de setores censitários, por situação do domicílio		
	Total	Urbano	Rural
Piauí	40	19	21
Ceará	42	28	14
Rio Grande do Norte	40	25	15
Paraíba	74	43	31
Pernambuco	45	29	16
Alagoas	43	26	17
Sergipe	72	43	29
Bahia	101	58	43
Minas Gerais	40	29	11
Espírito Santo	39	27	12
Rio de Janeiro	85	79	6
São Paulo	75	67	8
Paraná	40	27	13
Santa Catarina	65	46	19
Rio Grande do Sul	87	60	27
Mato Grosso do Sul	46	32	14
Mato Grosso	70	47	23
Goiás	40	29	11
Distrito Federal	40	36	4

Fonte: IBGE, Pesquisa de Avaliação da Cobertura da Coleta.
(1) Exclusive a área rural.

Em cada um dos setores escolhidos, foram selecionados domicílios particulares, com o intuito de calcular os indicadores de omissão de pessoas, inclusive por faixa etária, residentes em domicílios particulares ocupados, considerados os mesmos no censo e na PA.

Para a consecução desse objetivo, a escolha da fração amostral a ser aplicada de forma sistemática para seleção dos domicílios resultou de testes com valores iguais a 10%, 20% e 25%. De posse dos resultados, verificou-se que correspondiam a um nível de precisão adequado e que havia uma grande semelhança entre os coeficientes de variação para as três frações testadas, devido à reduzida variabilidade existente entre os domicílios. Assim sendo, ficou decidido que seria utilizada a fração de segundo estágio igual a 10%. A Tabela 10.2 fornece os resultados correspondentes.

Tabela 10.2 - Número de domicílios particulares ocupados e de pessoas residentes, por estágios da amostra da Pesquisa de Avaliação, segundo a situação do domicílio e Unidades da Federação - 2000

(continua)

Situação do domicílio e Unidades da Federação	Primeiro estágio		Segundo estágio	
	Domicílios	Pessoas	Domicílios (1)	Pessoas (2)
Brasil	**301 230**	**1 168 494**	**25 517**	**106 514**
Urbana	249 801	955 049	21 182	87 517
Rural	51 429	213 445	4 335	18 997
Rondônia	7 111	26 735	581	2 355
Acre	8 328	32 555	656	2 847
Amazonas	8 917	41 192	730	3 803
Roraima	5 416	22 455	457	2 107

Tabela 10.2 - Número de domicílios particulares ocupados e de pessoas residentes, por estágios da amostra da Pesquisa de Avaliação, segundo a situação do domicílio e Unidades da Federação - 2000

(conclusão)

Situação do domicílio e Unidades da Federação	Primeiro estágio		Segundo estágio	
	Domicílios	Pessoas	Domicílios (1)	Pessoas (2)
Pará	10 593	47 614	934	4 627
Amapá	6 451	30 057	522	2 711
Tocantins	12 081	48 333	951	4 149
Maranhão	8 493	37 962	710	3 408
Piauí	7 565	32 460	674	3 155
Ceará	8 573	35 656	753	3 358
Rio Grande do Norte	10 569	42 831	897	3 975
Paraíba	14 412	58 179	1 254	5 365
Pernambuco	11 369	43 846	917	3 820
Alagoas	10 863	46 331	932	4 296
Sergipe	14 072	56 844	1 242	5 318
Bahia	19 427	78 096	1 622	6 956
Minas Gerais	7 985	29 675	681	2 688
Espírito Santo	9 569	34 554	750	2 903
Rio de Janeiro	17 823	60 499	1 496	5 572
São Paulo	15 892	55 544	1 379	5 103
Paraná	8 908	31 732	771	2 911
Santa Catarina	15 497	54 829	1 400	5 146
Rio Grande do Sul	16 352	55 032	1 448	5 181
Mato Grosso do Sul	10 879	39 276	949	3 681
Mato Grosso	15 187	57 089	1 197	4 907
Goiás	9 923	34 930	857	3 127
Distrito Federal	8 975	34 188	757	3 045

Fonte: IBGE, Pesquisa de Avaliação da Cobertura da Coleta.
(1) Domicílios selecionados. (2) Pessoas confrontadas.

Para mais esclarecimentos sobre o dimensionamento da amostra da PA, consultar fontes de informação complementares na lista de Referências, ao final do capítulo.

10.3 Instrumentos de coleta

Boa parte dos instrumentos de coleta utilizados na PA são bastante semelhantes aos utilizados pelo censo. É apresentada aqui uma descrição geral desses documentos, dando ênfase às suas principais características. Para um conhecimento em maior detalhe, consultar os anexos desse capítuplo ou do CD-Rom que acompanha a metodologia.

PA 1.01 – Questionário da Amostra – Foi aplicado nos domicílios particulares ocupados selecionados para a amostra da PA. Contem dados relativos ao domicílio e, sobretudo, informações para as pessoas moradoras. Na sua primeira página são apresentados:

- os itens de identificação, o endereço do domicílio e os totais de pessoas pesquisadas e de questionários utilizados;

- a caracterização do domicílio quanto à espécie e tipo;

- a lista nominal, o sexo e a idade dos moradores, na data de referência do censo, que morreram ou mudaram-se até a data de referência da PA, bem como a data de sua saída do domicílio; e

- a lista nominal dos moradores, na data de referência da PA, e o ano em que passaram a residir no domicílio.

No restante do questionário, páginas 2 a 4, existe espaço para o registro de informações para até 9 moradores; foram investigados, para cada um, o nome, o sexo, a relação com a pessoa responsável pelo domicílio, o mês e ano de nascimento, a idade presumida, se ali residia na data de referência do censo e o mês e o ano nos quais passou a residir no domicílio.

PA 1.06 – Caderneta do Setor – Documento que contem o mapa e a descrição do setor, possuindo campos destinados à totalização dos registros, aos lançamentos das datas de início e fim da coleta, servindo também para o acondicionamento das Folhas de Coleta utilizadas.

PA 1.07 – Folha de Coleta – Destinava-se ao registro de todas as unidades residenciais e não-residenciais existentes no setor, bem como de seus endereços, da espécie de cada unidade, da numeração dos domicílios particulares para a seleção da amostra, e do nome da pessoa responsável pelo domicílio e do cônjuge, se houvesse. Contem, também, espaços destinados à totalização de questionários preenchidos e de pessoas recenseadas e à aplicação de códigos de batimento.

PA 1.09 – Manual do Recenseador – Documento que apresentava a finalidade da pesquisa, os conceitos, as definições e os procedimentos que deviam ser utilizados durante a coleta, bem como as instruções sobre o preenchimento dos demais instrumentos.

PA 1.11 – Manual de Coordenação, Supervisão e de Empastamento – Esse documento destinou-se aos Coordenadores da PA e aos Supervisores. Contem as normas e rotinas necessárias ao acompanhamento da coleta, à recepção, ao controle e ao empastamento e remessa do material coletado.

PA 1.18 e PA 1.19 – Folhas de Encaminhamento do Setor – Posto / Agência e Agência/UR – Destinaram-se à remessa de todo o material da coleta dos setores; dos postos de coleta às agências e destas para a sede da Unidade Regional.

10.4 Treinamento das equipes

O treinamento das equipes da coleta começou no Rio de Janeiro onde, durante três dias, capacitou técnicos pertencentes à DPE, que vieram a fazer parte do primeiro grupo de instrutores para a coleta da pesquisa. Mais tarde, ocorreu a instrução para os Coordenadores Estaduais da PA e técnicos dos quadros das Unidades Regionais, em quatro pólos de treinamento. A Tabela 10.3, a seguir, apresenta os quantitativos de treinandos em cada pólo de treinamento.

Tabela 10.3 - Número de treinandos, técnicos da Unidades Regionais e coordenadores, segundo pólos de treinamento - 2000

Pólos de treinamento	Número de treinandos
Total	**51**
Pará	13
Ceará	15
São Paulo	13
Goiás	10

Fonte: IBGE, Pesquisa de Avaliação da Cobertura da Coleta.

Na definição dos pólos foi levada em conta a proximidade geográfica das Unidades da Federação, resultando em:

- Pólo 1: Rondônia, Acre, Amazonas, Roraima, Pará, Amapá, Tocantins, Maranhão e Piauí;

- Pólo 2: Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe e Bahia;

- Pólo 3: Espírito Santo, Rio de Janeiro, São Paulo, Paraná, Santa Catarina e Rio Grande do Sul; e

- Pólo 4: Minas Gerais, Mato Grosso do Sul, Mato Grosso, Goiás e Distrito Federal.

Dando continuidade ao processo de capacitação das equipes de coleta da PA, ocorreram, em cada Unidade da Federação, mais dois repasses de instrução para os técnicos exclusivamente contratados para o trabalho de campo da pesquisa, ou seja:

- treinamento, em três dias, para os Supervisores, a cargo dos Coordenadores da PA e de alguns técnicos dos quadros das Unidades Regionais; e

- na seqüência, instrução, também em três dias, para os Recenseadores, a cargo da mesma equipe de instrutores, acrescida de alguns Supervisores indicados.

Os números de treinandos da PA são informados na Tabela 10.4 a seguir:

Tabela 10.4 - Número de treinandos, recenseadores e supervisores, segundo as Unidades da Federação - 2000

(continua)

Unidades da Federação	Número de treinandos	
	Recenseadores	Supervisores
Brasil	**1 354**	**319**
Rondônia	28	7
Acre	26	6
Amazonas	40	12
Roraima	26	6
Pará	40	10
Amapá	26	6

Tabela 10.4 - Número de treinandos, recenseadores e supervisores, segundo as Unidades da Federação - 2000

(conclusão)

Unidades da Federação	Número de treinandos	
	Recenseadores	Supervisores
Tocantins	40	9
Maranhão	44	11
Piauí	40	11
Ceará	42	11
Rio Grande do Norte	40	11
Paraíba	74	18
Pernambuco	45	10
Alagoas	43	10
Sergipe	72	16
Bahia	101	23
Minas Gerais	40	9
Espírito Santo	39	9
Rio de Janeiro	85	20
São Paulo	75	19
Paraná	40	9
Santa Catarina	65	14
Rio Grande do Sul	87	18
Mato Grosso do Sul	46	10
Mato Grosso	70	16
Goiás	40	9
Distrito Federal	40	9

Fonte: IBGE, Pesquisa de Avaliação da Cobertura da Coleta.

Posteriormente, ocorreu o treinamento para as fases de Batimento e Reconciliação das Informações coletadas pela PA. Como aconteceu na coleta dos dados, inicialmente, durante cinco dias, foram treinados 13 técnicos da DPE que, em conjunto com seus instrutores, formaram equipes para o repasse da instrução aos técnicos das Unidades Regionais. Para isso, foram constituídos oito pólos de treinamento, conforme mostra o Tabela 10.5.

Tabela 10.5 - Número de treinandos, segundo pólos de treinamento - 2000

Pólos de treinamento	Número de treinandos (1)
Total	**169**
Amazonas	22
Ceará	22
Bahia	20
Paraíba	19
Rio de Janeiro	21
São Paulo	21
Rio Grande do Sul	20
Goiás	24

Fonte: IBGE, Pesquisa de Avaliação da Cobertura da Coleta.

10.5 Coleta dos dados

A coleta da Pesquisa de Avaliação foi realizada em cada um dos setores, previamente selecionados, após o encerramento de sua coleta e supervisão do Censo, bem como da respectiva liberação de pagamento ao Recenseador. A data de referência da PA foi o dia da visita ao domicílio.

Os quantitativos de Supervisores em cada Unidade da Federação foram definidos levando-se em conta, além da carga de trabalho, a proximidade espacial dos municípios que continham setores selecionados para a pesquisa – cada Supervisor ficou responsável, em média, por quatro recenseadores.

Tendo em vista que os setores a serem coletados na PA foram divulgados às Unidades Regionais paulatinamente, foi enviado com antecedência, para cada unidade, um arquivo contendo o número de setores selecionados por município e subárea. A efetiva identificação dos setores para o início da coleta, foi feita a partir dos primeiros dias do mês de outubro, inicialmente através do correio eletrônico e, posteriormente, através do Sistema de Acompanhamento da Coleta, Batimento e Reconciliação.

A seguir, faz-se uma descrição sumária das atribuições de cada técnico envolvido na atividade de coleta da PA, sendo que grande parte desse trabalho foi desenvolvida de maneira bastante semelhante ao que foi feito no censo.

As tarefas a cargo do Coordenador Estadual da PA começaram antes do início do trabalho de campo da pesquisa e tiveram continuidade durante todo o período da coleta. De maneira geral, suas atribuições eram:

- organizar e administrar toda a coleta na Unidade Regional;

- organizar e participar do treinamento dos Supervisores e Recenseadores;

- alocar os Supervisores em suas respectivas áreas de trabalho;

- providenciar o material necessário para a execução da coleta, bem como orientar sua distribuição;

- orientar e providenciar a montagem das Cadernetas dos Setores selecionados;

- orientar os Supervisores quanto à distribuição dos setores aos Recenseadores;

- estabelecer um roteiro de visitas às agências e aos postos de coleta, para supervisão dos trabalhos;

- atualizar, semanalmente, as informações do Sistema de Acompanhamento da Coleta, Batimento e Reconciliação; e

- organizar a recepção e conferência do material coletado, remetido pelas agências de coleta à sede da Unidade Regional.

O trabalho dos Supervisores era também bastante parecido com o desenvolvido durante o censo, sendo a diferença determinada, sobretudo, pela aplicação de uma rotina mais simplificada do plano de supervisão para o material coletado. Assim, ao Supervisor cabia, antes do início da coleta:

- localizar e percorrer os setores sob sua responsabilidade, verificando se as descrições de seus limites e o mapa refletiam a situação em campo; e

- orientar os recenseadores quanto à perfeita identificação do setor.

Durante o período de coleta, cabia ao supervisor:

- orientar os Recenseadores, sobretudo durante a primeira semana de trabalho, quanto ao preenchimento dos instrumentos de coleta e ao processo de seleção dos domicílios da amostra; nessa ocasião, deveriam realizar reentrevistas em, pelo menos, cinco domicílios em cada setor.

- organizar a entrega semanal do material coletado, marcando dia e hora para que os Recenseadores comparecessem ao posto de coleta;

- verificar criteriosamente o preenchimento dos questionários e das Folhas de Coleta;

- selecionar, ao menos, um trecho de cada setor e fazer o seu percurso, verificando a correção da listagem das unidades, principalmente no que se refere à omissão de domicílios; e

- informar semanalmente sobre o andamento dos trabalhos de coleta, identificando os setores não iniciados, em trabalho, interrompidos e liberados.

Por fim, encerrada a coleta do setor, era ainda tarefa do Supervisor:

- revisar os lançamentos da Caderneta do Setor;

- conferir as quantidades utilizadas de todo o material de coleta; e

- fazer a separação e o empastamento do material, providenciando sua remessa à agência de coleta.

Na PA, o trabalho de listagem ou cobertura do setor, a cargo do Recenseador, foi realizado exatamente como no censo. Nesse aspecto, os conceitos, definições e os procedimentos adotados, obrigatoriamente, foram os mesmos nas duas pesquisas, em função de que o objetivo da PA foi avaliar a cobertura da coleta do censo.

Já a tarefa de visitar domicílios para coleta das informações da PA guardou semelhança com a correspondente atividade censitária somente no que se refere à atitude do Recenseador, às técnicas de abordagem do informante e à condução da entrevista. No mais, as diferenças eram bastante acentuadas, tendo em vista os conteúdos distintos dos questionários das duas pesquisas, e também o fato da PA investigar, somente, uma amostra dos domicílios ocupados no setor.

A atualização periódica do Sistema de Acompanhamento da Coleta, Batimento e Reconciliação permitiu ao coordenador da pesquisa e à equipe responsável na DPE fazerem um acompanhamento contínuo da evolução do trabalho de campo. Dessa forma, diariamente, era possível, conhecendo o número de setores não iniciados, com a coleta em andamento, interrompidos e liberados, avaliar o avanço da coleta, tomando as medidas gerenciais cabíveis. O sistema fornecia para cada setor, a data de início da coleta e, quando da sua liberação pelo Supervisor, ficavam também disponíveis as informações constantes da capa da caderneta, ou seja, os totais de unidades visitadas e pessoas recenseadas, bem como o número de domicílios particulares e de domicílios particulares selecionados, segundo a espécie.

10.6 Batimento e reconciliação dos dados

As fases de Batimento e Reconciliação das informações da PA ficaram sob os cuidados diretos do coordenador da pesquisa e de técnicos do quadro do IBGE, com dedicação exclusiva a essas tarefas que foram desenvolvidas nas sedes das Unidades Regionais, menos, é claro as visitas aos domicílios.

Como se sabe, o êxito da PA está intimamente ligado ao tempo, que deva ser o menor possível, entre a sua coleta e a do censo. No entanto, é importante acrescentar que essa condição de período mínimo também se verifique entre o término da coleta da PA e o início da fase de batimento e desta para a de Reconciliação das Informações.

A implementação da fase do Batimento representou a primeira parte da apuração da pesquisa, posto que ao final desses trabalhos foi obtida, sob a forma de códigos, a consolidação das ocorrências encontradas entre as coletas do censo e da PA.

A fase da Reconciliação ou Checagem das informações constituiu o complemento do trabalho de batimento, tão relevante quanto a anterior, no sentido de que em algumas das unidades codificadas podiam ocorrer situações de dúvida, necessitando-se assim de esclarecimentos adicionais através do retorno ao domicílio.

As bases para a realização dessas tarefas foram, por um lado, o Manual de Batimento, Reconciliação e Crítica – PA 1.10 – e, por outro, o Sistema de Acompanhamento da Coleta, Batimento e Reconciliação. O Manual orientou detalhadamente cada passo das tarefas, enquanto o sistema deu o tratamento necessário às informações, gerenciando a entrada dos dados, fazendo as críticas de validade e consistência e oferecendo o necessário ao acompanhamento dessas duas importantes fases do trabalho da PA.

Por fim, é importante lembrar que todo o material de coleta do censo referente aos setores selecionados para a PA permaneceu nas Unidades Regionais, de modo a ser possível realizar as tarefas relativas ao Batimento e Reconciliação das Informações.

10.6.1 Batimento dos dados

O batimento das informações coletadas pelas duas pesquisas foi realizado logo após o término da coleta e supervisão de cada setor na PA. Inicialmente, foi realizada uma comparação dos mapas e descrição dos limites do setor, constantes das Cadernetas do Setor do censo e da PA, tendo em vista assegurar que correspondiam à mesma área de coleta. Em seguida, fazia-se uma comparação entre os dados do resumo de cada uma delas, objetivando avaliá-los grosseiramente, identificando a ordem de grandeza das possíveis diferenças entre o número de domicílios ocupados e pessoas recenseadas, e se houve invasão de um setor vizinho ou falhas de cobertura.

A fase do Batimento das informações começava efetivamente com a comparação de todas as unidades residenciais ou não, registradas nas Folhas de Coleta do Censo e da PA. Para cada unidade comparada, era atribuído um código de batimento de três dígitos. O primeiro caracterizava a situação da unidade avaliada na PA; o segundo indicava a sua situação com respeito ao censo; e o último fazia a caracterização simultânea nas duas pesquisas. Por exemplo, o código 112 correspondia a uma unidade classificada como domicílio ocupado no censo e na PA, onde houve pessoas recenseadas indevidamente. No setor em que unidades visitadas pela PA não puderam ser identificadas no correspondente setor do censo, foi investigada a possibilidade dessas unidades terem sido listadas em setores limítrofes.

As atividades de batimento das Folhas de Coleta se encerravam com o lançamento na PA 1.12 – Tela 1 – e desta para o Sistema de Acompanhamento da Coleta, Batimento e Reconciliação, de todos os códigos atribuídos. Também era preenchida o PA 1.16 - Boletim1, onde se identificavam somente os domicílios que apresentavam códigos divergentes – 113, 121, 131, 141, 161, 171, 101, 470 e 710 –, que seriam trabalhados na fase de reconciliação.

Uma vez concluído o batimento das Folhas de Coleta, teve início o batimento dos questionários. Para isso, além das folhas anteriormente utilizadas, foi necessário, é claro, ter-se acesso aos questionários preenchidos no censo. Foram objeto de comparação os questionários dos domicílios particulares ocupados e selecionados para a amostra da PA, que receberam os seguintes códigos:

- 111 – um domicílio particular ocupado no censo correspondia a um domicílio particular ocupado na PA;

- 115 – um domicílio particular ocupado no censo correspondia a dois ou mais domicílios particulares na PA; e

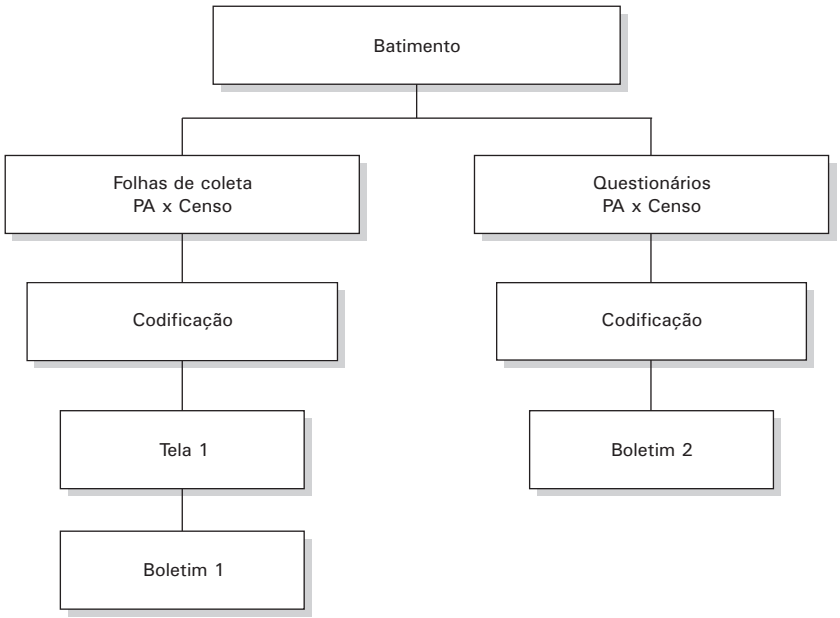
- 117 – dois ou mais domicílios particulares ocupados no censo correspondia a um único domicílio particular ocupado na PA.

Nesse trabalho também foram aplicados códigos às pessoas, de modo que retratassem a situação de moradia de cada uma delas na PA e no censo. Assim, podia-se ter, por exemplo: recenseamento indevido – código 70; omissão de pessoa no censo – código 00.

A identificação de todos os domicílios e dos moradores, correspondentes às situações de omissão de pessoas, no censo ou na PA, bem como as dúvidas que não permitiram a codificação, foram lançadas no PA 1.17 – Boletim 2 – para serem investigadas em campo, durante a fase de reconciliação.

As tarefas do batimento podem ser resumidas esquematicamente da seguinte forma:

Figura 10.1 - Rotina do Batimento



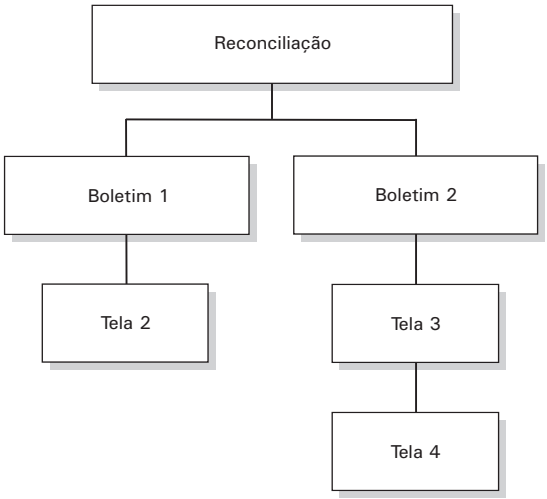
10.6.2 Reconciliação dos dados

Para a implementação dessa fase foram utilizados os mesmos técnicos que realizaram o batimento, de modo que, sendo conhecedores de cada situação divergente ou duvidosa, pudessem solucioná-la com maior facilidade.

A reconciliação consistiu na checagem, no campo, das informações dos domicílios e das pessoas, visando ao esclarecimento das divergências ou dúvidas de codificação encontradas na fase de batimento.

Nessa ocasião, alguns códigos registrados no Boletim 1, relativos ao batimento das Folhas de Coleta, e no Boletim 2, referentes ao batimento dos questionários, podiam ser confirmados ou revistos. Os dados, alterados ou não, no Boletim 1 foram registrados na PA 1.13 – Tela 2, enquanto os do Boletim 2 tiveram sua consolidação na PA 1.14 – Tela 3 e na PA 1.15 – Tela 4. A rotina de trabalho da reconciliação pode ser esquematizada como:

Figura 10.2 - Rotina de Reconciliação



Por fim, as informações da PA 1.14 e PA 1.15 foram carregadas no Sistema de Acompanhamento da Coleta, Batimento e Reconciliação, que implementou as funções de crítica, disponibilizando as informações necessárias ao cálculo das estimativas dos diversos indicadores de cobertura da PA.

10.7 Métodos de estimação

Como assinalado no início deste capítulo, foi utilizado na PA do Censo 2000 o método *Dual System Estimation,* para estimação dos indicadores de cobertura da coleta.

O modelo estatístico utilizado foi originalmente desenvolvido para estimar o tamanho de populações fechadas em estudos biométricos. Diz-se população fechada aquela que permanece praticamente inalterada durante a condução do estudo. Baseia-se na técnica de estimação *capture-recapture* desenvolvida para estimar o tamanho de populações de peixes ou animais (FAY; PASSEL; ROBINSON, 1988).

Esse método requer duas diferentes fontes de dados, independentes, da população a ser medida, e exige o confronto das informações obtidas por essas fontes.

O quadro, a seguir, apresenta o resultado do confronto dos dados das duas pesquisas:

Quadro 10.1 - Resumo do confronto dos dados das pesquisas - 2000

Segunda captura	Primeira captura		
	Total	Unidades incluídas	Unidades não-incluídas
Total	**a+b**	**a+c**	**b+d**
Unidades incluídas	a+b	a	b
Unidades não-incluídas	c+d	c	d

onde:

a é o número de unidades incluídas em ambas as pesquisas;

b é o número de unidades incluídas apenas na segunda pesquisa;

c é o número de unidades incluídas apenas na primeira pesquisa;

d é o número de unidades não-incluídas nas duas pesquisas; e

t é o total de unidades da população.

Considerando-se que em cada setor selecionado a primeira pesquisa realizada seja o censo e a segunda a PA, pode-se deduzir dessa relação o estimador para o tamanho da população proveniente do método *capture-recapture*, também chamado estimador do Sistema Dual, estimador de Petersen ou de Sekar-Deming, que é dado por:

$$\hat{t} = \frac{(a+c)\,(\hat{a}+\hat{b})}{\hat{a}}$$

onde:

$(a+c)$ é a população apurada pelo censo;

$(\hat{a}+\hat{b})$ é o estimador do total da população apurado pela PA;

$\hat{a}$ é o estimador do total da população incluída nas duas pesquisas.

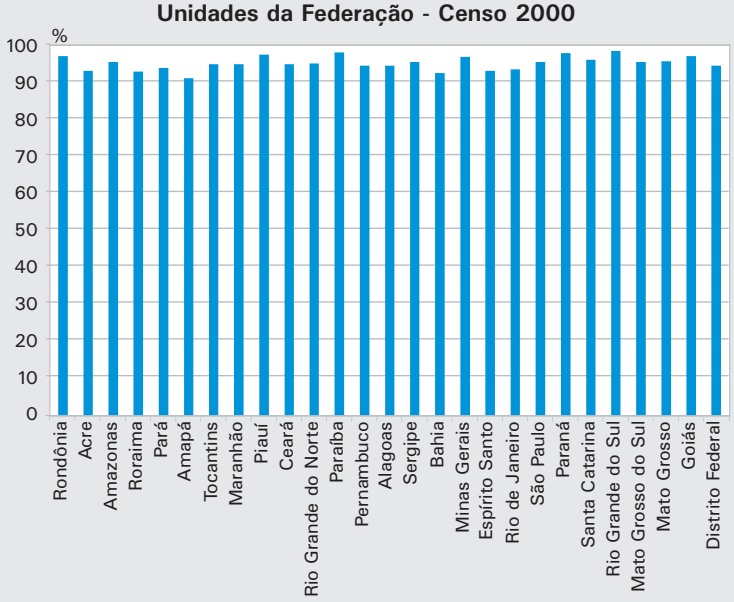
De uma forma geral, a taxa de omissão é dada por

$$\hat{R} = \frac{\hat{b}}{(\hat{a}+\hat{b})}$$

sendo a taxa de cobertura é igual a $\hat{C} = 1 - \frac{\hat{b}}{(\hat{a}+\hat{b})}$

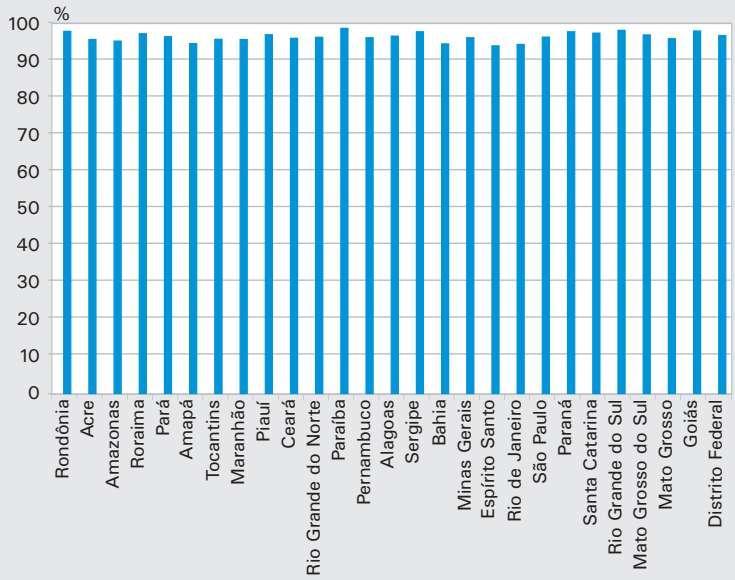
Em seguida são apresentados gráficos de duas das taxas de cobertura da coleta do Censo 2000, calculadas com base na PA.

Gráfico 10.1 - Níveis de cobertura referentes às pessoas residentes em domicílios particulares ocupados, para as Unidades da Federação - Censo 2000



Fonte: IBGE, Pesquisa de Avaliação da Cobertura da Coleta.

Gráfico 10.2 - Níveis de cobertura referentes aos domicílios particulares ocupados, por Unidade da Federação – Censo 2000



Fonte: IBGE, Pesquisa de Avaliação da Cobertura da Coleta.

O resultado completo da pesquisa, apresentando a definição de todas as taxas calculadas, bem como os comentários acerca de seus valores, pode ser encontrado em Oliveira e outros (2003).

Referências

CENSO demográfico 2000: pesquisa de avaliação da cobertura da coleta: manual de batimento, reconciliação e crítica – PA 1.10. Rio de Janeiro: IBGE, 2000.

_____: manual de coordenação, supervisão e de empastamento – PA 1.11. Rio de Janeiro: IBGE, 2000.

_____: manual do recenseador – PA 1.09. Rio de Janeiro: IBGE, 2000.

COSTA, L. N. *Estudos e pesquisas de avaliação de censos demográficos* - 1970 a 1990. Rio de Janeiro: IBGE, 1990. (Textos para discussão, n. 34).

FAY, R. E.; PASSEL, J. S.; ROBINSON, J. G. *1980 census of population and housing:* evaluation and research reports: the coverage of population in the 1980 census. Washington, D.C.: U.S. Bureau of the Census, 1988. 123 p.

OLIVEIRA, L. C. de S. et al. *Censo demográfico 2000:* resultados da pesquisa de avaliação da cobertura da coleta. Rio de Janeiro: IBGE, 2003. 142 p. (Textos para discussão, n. 9).

OLIVEIRA, L. C. de S.; FREITAS, M. P. S.; BIANCHINI, Z. M. *Pesquisa de avaliação da cobertura da coleta do censo demográfico 2000:* definição do desenho amostral. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1999.

RELATÓRIO da pesquisa de avaliação do censo experimental 1999. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2000. 23 p.

SELTZER, W.; ADLAKHA, A. *On the effect of errors in the application of the Chandrasekar-Deming Thecnique.* New York: University of North Carolina, 1974. 13 p. (Laboratories for population statistics reprint series, n. 14).

Anexos

PA 1.01 Questionário da Amostra da Pesquisa de Avaliação da Cobertura da Coleta do Censo Demográfico 2000

PA 1.06 Caderneta do Setor da Pesquisa de Avaliação da Cobertura da Coleta do Censo Demográfico 2000

PA 1.07 Folha de Coleta da Pesquisa de Avaliação da Cobertura da Coleta do Censo Demográfico 2000

11 Captura dos dados

A tecnologia que foi utilizada para captura de dados permitiu o reconhecimento de campos com código de barras, marcas e caracteres – numéricos e alfanuméricos – a partir da digitalização dos questionários. Isso efetivamente agilizou a atividade de entrada dos dados coletados no Censo.

Para o reconhecimento de números e textos manuscritos, foi exigido um nível mínimo na qualidade da escrita por parte do Recenseador. A solução adotada garantiu que a informação pudesse ser interpretada dos questionários no estado em que estivessem. Assim, foi garantido o ingresso de todos os questionários no processamento, bem como a obtenção de todas as imagens dos mesmos.

No processo de reconhecimento, com interpretação de marcas e caracteres, sempre existe um erro associado. Como tal situação é inerente ao processo, foi necessário manter as taxas de erro em níveis aceitáveis. Assim, foi implementado um Sistema de Controle de Qualidade do Reconhecimento, para monitorar essas taxas durante toda a execução do processo de captura de dados, onde se estabeleceu um nível de qualidade mínimo, necessário para cada tipo de campo do questionário.

Os casos em que não houve reconhecimento foram resolvidos numa etapa posterior, denominada verificação. Cabe registrar que, antes dos dados serem transmitidos para o equipamento central, passaram por uma etapa de crítica quantitativa.

O serviço de desenvolvimento do sistema para o processamento das imagens e captura dos dados foi contratado a uma empresa especialista no assunto. A atuação dos técnicos da Instituição concentrou-se não apenas na coordenação, mas também

na especificação do desenvolvimento da solução. A montagem da linha de produção, bem como a operacionalização de todo o processo, ficaram, também, a cargo dos técnicos da Instituição.

11.1 Dados dos questionários

A captura dos dados foi efetuada em cinco Centros de Captura de Dados – CCDs –, localizados nas cidades do Rio de Janeiro/RJ, Campinas/SP, Goiânia/GO, Campina Grande/PB e Curitiba/PR. O sistema utilizado para essa captura foi baseado na digitalização dos documentos por meio de *scanners*, além de contar com o reconhecimento, através de *software*, dos códigos de barras, dos campos de marcas (OMR), de caracteres pré-impressos (OCR) e dos campos com caracteres manuscritos (ICR).

A captura dos dados foi feita de forma simultânea para os Questionários Básico (CD 1.01) e da Amostra (CD 1.02). Nessa etapa, o controle foi feito por modelo de questionário e por setor censitário. O quadro abaixo mostra os Centros de Captura de Dados e as Unidades da Federação abrangidas.

Quadro 11.1 - Captura dos dados coletados nas Unidades da Federação - 2000

Centros de Captura de Dados	Unidades da Federação
Rio de Janeiro/RJ	RO, AC, AM, RR, PA, AP, TO
Campinas/SP	SP, ES, MS
Goiânia/GO	MG, GO, MT, DF
Curitiba/PR	PR, SC, RS, RJ
Campina Grande/PB	MA, PI, CE, RN, PB, PE, AL, SE, BA

Cada Centro de Captura de Dados foi gerenciado de forma independente, havendo o controle de todo o processo de captura, com vistas a garantir o prazo estabelecido para essa etapa.

Com base em experimentos realizados e na captura dos dados do Censo Experimental, foram planejados 15 módulos de produção, contendo cada módulo dois *scanners*. Os CCDs de Campinas, Campina Grande e Curitiba foram dotados de quatro módulos de produção. O CCD de Goiânia trabalhou com dois módulos, enquanto o do Rio de Janeiro contava com um módulo de produção.

Para os técnicos contratados para trabalhar nos CCDs (auxiliares censitários de apuração e auxiliares censitários de informática) foram realizados treinamentos técnico-operacionais, tendo como objetivo o conhecimento da missão institucional do IBGE e das tarefas inerentes às suas funções. Para os auxiliares de informática, foi planejado um treinamento de capacitação e seleção para as funções de Supervisores de Módulo e de Turno. Esse treinamento foi efetuado através de manuais, exercícios e dinâmica de grupo e ministrado diretamente nos CCDs por técnicos do CCD do Rio de Janeiro e da Escola Nacional de Ciências Estatísticas – ENCE. Os módulos tinham como objetivo validar todo o controle de captura das informações, desde a recepção até o tratamento de problemas ocorridos na digitalização dos questionários.

A Tabela 11.1 a seguir apresenta os contingentes utilizados nos CCDs e suas diversas funções:

Tabela 11.1 - Efetivo de pessoal dos Centros de Captura de Dados - CCDs - 2000

Funções	Efetivo de pessoal dos Centros de Captura de Dados - CCDs					
	Total	Rio de Janeiro	Goiás	Paraíba	Paraná	São Paulo
Total	**1 321**	**104**	**218**	**329**	**335**	**335**
Recepção	42	5	7	10	10	10
Preparação	112	10	18	28	28	28
Operador de *scanner*	186	12	30	48	48	48
Verificação/crítica	774	48	126	196	202	202
Operação CPD	44	6	8	10	10	10
Apoio administrativo	35	5	6	8	8	8
Apoio operacional	44	6	8	10	10	10
Coordenadores	24	4	5	5	5	5
Supervisor de módulo	40	4	6	10	10	10
Supervisor de CPD	10	2	2	2	2	2
Supervisor de turno	10	2	2	2	2	2

Fontes: IBGE, Coordenação de Acompanhamento e Controle Operacional dos Censos e Diretoria de Informática.

11.1.1 Recepção

Ainda na etapa de recepção, os questionários eram organizados em ordem crescente pelo quesito 1.01. Na falta de numeração dos questionários, estes deveriam ser numerados com a orientação e autorização do Coordenador Técnico, conforme as instruções de preenchimento no campo.

Na recepção, os questionários vinham acondicionados em caixas, separados por tipos: básico e amostra. Após retirados das caixas, eram conferidos, contados e registrados com base nas informações do setor. Esse procedimento de conferência tratava o material enviado pelas unidades regionais e visava, especificamente, prevenir o extravio ou a inclusão indevida de questionários no setor. Inicialmente, as quantidades dos lotes de questionários de cada setor eram confrontadas com os totais de questionários coletados e informados no Sistema de Indicadores Gerenciais da Coleta - SIGC.

Para tratar as possíveis divergências entre esses números, foram estabelecidos parâmetros de controle, que determinavam limites admissíveis de variação, tendo em conta a natureza do trabalho realizado na coleta de dados. No caso do CD 1.01, os parâmetros levaram em consideração os números de questionários do setor, ou seja: até 50; de 51 até 250; de 251 até 500; e acima de 500. De acordo com essa estrutura, os parâmetros eram, respectivamente, 5%; 3%; 2% e 1%. Já para o CD 1.02, a maior diferença admitida era de 10%, independentemente do número de questionários do setor.

Caso as diferenças superassem esses limites, o lote era rejeitado, sendo consultada a Divisão de Pesquisa – DIPEQ da Unidade Regional responsável pela coleta do setor, para que tomasse as providências cabíveis, com vistas ao esclarecimento da divergência. Resolvidas todas as pendências, gravava-se no Sistema de Indicadores Gerenciais da Produção – SIGP o quantitativo de questionários com valores do campo 1.01 – Número do Questionário – diferentes, o que representava, naquele momento, o número corrigido de unidades entrevistadas no setor. Entenda-se como unidade entrevistada o domicílio particular ou

ainda a família ou a pessoa só em domicílio coletivo, visitados pelo Censo e onde foi possível obter a entrevista.

11.1.2 Preparo e digitalização do material

Nessa etapa, as pastas numeradas e os questionários organizados deveriam ser preparados para a digitalização. Para os questionários da amostra, utilizava-se a máquina saca-grampo para a retirada dos mesmos e, em todo o material verificado, levavam-se em conta os seguintes itens: clipes, ausência de código de barras, riscos, preenchimentos com caneta, rasgões, manchas e umidade. Para as diversas situações encontradas, era necessário um tratamento específico: os questionários rasgados eram restaurados com fita durex; os questionários amassados e úmidos eram passados a ferro de passar roupa; e os questionários sem o código de barras eram numerados manualmente, a partir da relação de códigos não utilizados na impressão dos questionários. Após esses procedimentos de preparo, o material estava pronto para ser digitalizado.

Ainda na etapa de preparação, as pastas já tratadas eram registradas no sistema para geração da capa de lote que, após digitalização identificava para o sistema as seguintes informações: modelo do questionário, número da caixa, número de seqüência da caixa, quantidade de questionários na caixa e, quando necessário, quantidade de folhas substituídas e os códigos de barra das folhas originais e substitutas.

Já com as capas de lote, as pastas eram encaminhadas para a digitalização, onde o operador, através da tela do computador, comandava o início da operação, informando o tipo de questionário a ser digitalizado. Essa tela também possuía as funções de preparação, captura e tratamento das imagens, bem como a configuração do *scanner*.

Era exigida atenção redobrada por parte dos operadores, sendo este procedimento realizado por três técnicos, adotando-se um rodízio entre eles: um operador organizava as pastas do carrinho para o *scanner* e acondicionava os questionários na bandeja de alimentação de material do equipamento; outro operador acionava o sistema, além de fornecer o número do lote para ser anotado pelo primeiro operador; e o terceiro operador verificava e acompanhava as imagens dos questionários digitalizados. A partir da digitalização, era formado o arquivo de imagens do lote e posteriormente enviado para o servidor de imagens.

11.1.3 Reconhecimento ótico

O processo de reconhecimento ótico envolveu o tratamento de quatro tipos de campos de informação: código de barras, marcas, alfabéticos e numéricos. Em tal processo, as imagens das páginas digitalizadas, passaram por identificação, enquanto as imagens dos campos selecionados eram fragmentadas. Os campos, por sua vez, foram segmentados e submetidos ao *software* específico de reconhecimento do código de barras, marcas, números e caracteres. Os campos reconhecidos, de acordo com um nível de confiabilidade previamente estabelecido, seguiam para o processo de arquivamento, e os não reconhecidos eram encaminhados para o processo de verificação.

O primeiro passo no reconhecimento foi a identificação da página do questionário para, em seguida, ser feito o reconhecimento dos respectivos campos.

Para os campos com códigos de barras, foi utilizado o *software* CLEQS, específico para essa situação. O nível de reconhecimento foi da ordem de quase 100%, só não sendo bem sucedido nos casos em que o código apresentava rasuras, riscos, ou estava borrado.

Os campos de marcas foram reconhecidos com a utilização do *software* AFPSPro, apresentando resultados bem satisfatórios, chegando à cifra de 99%. A presença de campos de marcas nos questionários é significativa, principalmente no Questionário da Amostra.

Os campos numéricos foram reconhecidos com a utilização de três *softwares* distintos: CGK, JustICR e Kadmos. Para cada número a ser reconhecido, o sistema obtinha três indicações (uma de cada *software*) e, num esquema de votação por maioria simples, determinava o valor a ser estabelecido. No caso de empate, utilizava a indicação do *software* JustICR, pois foi o que apresentou melhores resultados nas etapas de teste. Foram realizados vários testes, a partir de questionários preenchidos, com a finalidade de alimentar o sistema de autoaprendizado (próprio de cada *software*), nas diversas formas de grafia encontradas. Obteve-se uma média de reconhecimento da ordem de 92%.

Os campos alfabéticos também foram reconhecidos com o emprego dos mesmos *softwares* utilizados para os campos numéricos. Para cada letra a ser reconhecida, o sistema agia de forma semelhante ao descrito para os campos numéricos, havendo também o esquema de votação, sendo o *software* JustICR também utilizado para desempate. Para os campos alfabéticos, o sistema utilizava dicionários auxiliares, para complementar o reconhecimento. Esses dicionários foram obtidos no sistema de codificação automática e ajudavam o sistema a complementar uma palavra, no caso em que uma ou, no máximo, duas letras não tivessem sido reconhecidas. A utilização era feita sempre que somente uma entrada no dicionário era encontrada e só para substituição das letras não-reconhecidas. O dicionário fornecia um contexto de palavras para cada quesito que estava sendo reconhecido, sendo, portanto, utilizados diferentes dicionários no processo de reconhecimento. Obteve-se uma média de reconhecimento da ordem de 77%.

11.1.4 Digitação dos campos com baixo nível de reconhecimento

Essa etapa, denominada verificação, tinha como objetivo visualizar as imagens dos campos que não tinham atingido o nível mínimo estabelecido para o reconhecimento e, portanto, foram considerados pelo sistema como não-reconhecidos. Os campos eram apresentados na tela da maneira como tinham sido reconhecidos, juntamente com a sua imagem. Os números e letras que não foram reconhecidos pelo sistema apareciam com o símbolo '*'. Cabia ao operador analisar as duas informações e digitar o valor correto. A instrução existente era para que o operador preservasse o conteúdo do questionário, digitando o que realmente nele estava escrito (o que aparecia na imagem), mesmo que houvesse erro de grafia ou engano por parte do Recenseador (por exemplo, ano de nascimento igual a 2997).

Os campos de marcas que apresentavam dupla marcação, por falha de reconhecimento ou por erro de preenchimento, eram apresentados ao operador para que fosse feita a correção ou a confirmação da dupla marcação; no caso em que o operador confirmasse a dupla marcação, o campo, automaticamente, ficava sem preenchimento.

11.1.5 Consolidação dos dados

Este processo verificava se todas as páginas dos questionários de um lote estavam presentes com as respectivas páginas de frente e verso. Os questionários foram identificados por um código de barras, que era igual para todas as páginas de um mesmo questionário, variando apenas os dois últimos dígitos que as identificavam.

As inconsistências eram enviadas para tratamento na consolidação de códigos de barras, que consistia em corrigir os códigos das páginas com problemas ou substituí-los com base numa relação de códigos. Posteriormente, todos os conjuntos válidos de questionários de um mesmo setor, que podiam pertencer a lotes diferentes (diferentes pastas físicas), vindos das etapas anteriores, geravam um arquivo para ser submetido às próximas etapas da captura de dados.

11.1.6 Controle de qualidade

É cada vez mais freqüente a criação de rotinas e metodologias para a melhoria da qualidade dos produtos nas mais diversas áreas, inclusive na produção estatística. O fator qualidade é considerado essencial para o bom andamento dos trabalhos e para a credibilidade junto aos usuários dos dados produzidos. Para um dado processo de produção, o controle estatístico é importante na identificação de fatores capazes de comprometer o nível de qualidade desejado. Muitas situações especiais, ocorridas durante a produção do dado, podem introduzir variabilidade acima da considerada normal, ocasionando não-conformidades, incompatíveis com os padrões aceitáveis de qualidade.

Para o Censo 2000, o controle estatístico de processos foi o mecanismo para validar a qualidade dos processos de reconhecimento automático e digitação – verificação dos dados coletados nos questionários. Além disso, o contrato assinado com a empresa contratada para fornecer o sistema de reconhecimento ótico previa a realização de controle de qualidade, sob responsabilidade do IBGE, e estabelecia os parâmetros de qualidade a serem respeitados.

Esta Seção apresenta as ferramentas utilizadas para o monitoramento da variabilidade e da estabilidade dos processos, não apenas relatando os problemas verificados, mas também apresentando as soluções propostas para cada questão. A Seção está estruturada em duas partes. Inicialmente, é apresentado o controle de qualidade do reconhecimento. Além da metodologia adotada, são relatados os problemas verificados, as estatísticas gerais e as medidas de correção tomadas. Por fim, a mesma análise é feita para o controle de qualidade do processo de verificação. De fato, a etapa chamada de verificação consistia em um processo de digitação dos dados dos questionários que não haviam sido reconhecidos automaticamente pelo sistema de reconhecimento ótico.

11.1.6.1 Reconhecimento ótico

O Controle de Qualidade do Reconhecimento teve como finalidade controlar a tarefa de transcrição dos dados coletados nos questionários do Censo para o meio de armazenamento magnético, de tal modo que as informações armazenadas fossem as mais fiéis possíveis às registradas nos questionários. Tal processo foi efetivado por meio da digitação dos dados de amostras de questionários selecionados para comparação com as informações obtidas pelo reconhecimento automático.

A execução do plano de controle de qualidade do reconhecimento foi realizada, de forma descentralizada, nos cinco Centros de Captura de Dados.

Metodologia

O Controle de Qualidade do Reconhecimento - CQR teve como base de consulta o sistema de banco de dados gerados no processo de reconhecimento ótico dos questionários do Censo Demográfico 2000. Em cada turno de trabalho, o sistema selecionava, diariamente, 43 questionários do tipo Básico (CD 1.01) e 27 do tipo Amostra (CD 1.02). Para maiores detalhes sobre esse processo, ver Albieri e outros (2001).

As variáveis utilizadas pelo controle de qualidade do reconhecimento foram as seguintes:

- módulo: identificação do módulo. Dependendo do CCD, foram de um até quatro módulos, a saber: Rio de Janeiro, 1, Campinas, 4, Curitiba, 4, Goiânia, 2 e Campina Grande, 4;
- data;
- turno: identificação do turno usado para a seleção da amostra. Foram definidos os seguintes horários: turno 1, de 0:00h às 5:00h, turno 2, de 8:00h às 13:00h e turno 3, de 16:00h às 21:00h;
- modelo: 1 (básico) e 2 (amostra);
- questionários: número de questionários selecionados para o CQR;
- campo: indicador do tipo de campo, podendo ser 1 (campo alfanumérico) e 2 (campo numérico);
- preenchidos: número total de campos preenchidos nos questionários selecionados para o CQR;
- reconhecidos: número total de campos reconhecidos pelo *software* de reconhecimento; e
- divergentes: número total de campos em que os valores digitados não corresponderam ao o que foi reconhecido automaticamente pelo sistema computacional.

Para a obtenção de uma taxa líquida de divergência, que refletisse a qualidade da atuação do sistema de reconhecimento ótico das informações, dois procedimentos foram necessários. Primeiro, a identificação dos tipos de erro causadores de divergência. Segundo, a classificação dos erros deveria ser iniciada apenas quando o último questionário selecionado passasse pela etapa da verificação. Esses procedimentos foram realizados por Supervisores de Controle de Qualidade em cada um dos centros de captura. Para iniciar a classificação, os supervisores priorizavam os lotes selecionados e, em seguida, um sistema de consulta ao banco de dados fornecia as imagens de campos divergentes para a classificação. Para a realização dessas atividades, todos os supervisores receberam treinamento no Rio de Janeiro, que visou, principalmente, homogeneizar os critérios de classificação. Além das orientações repassadas para a classificação dos erros, o treinamento também consistiu na apresentação de noções básicas sobre ferramentas de controle estatístico (como gráficos de controle) e sobre o programa estatístico usado na análise dos resultados.

Os erros causadores de divergência foram classificados da seguinte forma:

Erro de Reconhecimento: todos os casos em que a divergência era causada por falha do *software* de reconhecimento. Para essa classificação, o supervisor deveria ser capaz de identificar duas situações:

1) casos em que o *software* reconhecia errado, ainda que a imagem estivesse sem nenhum problema.

Figura 11.1 – Exemplo de erro de reconhecimento

Tem-se um caso de erro clássico do *software*, pois o carácter divergente é claramente um "U", em vez de um "V", como o que foi reconhecido.

2) casos em que o *software* errava por reconhecer excessivamente, ainda que a imagem do carácter em questão estivesse fraca, como o caso apresentado abaixo:

Figura 11.2 – Exemplo de erro de reconhecimento excessivo

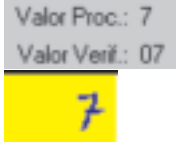
Em vez de enviar o carácter em questão para a verificação, o *software* reconhecia equivocadamente, daí a caracterização "erro excessivo".

Erro do Dicionário: Na verdade, era um outro tipo de erro de reconhecimento do *software*. Consistia na situação em que a atuação do dicionário de termos auxiliares do sistema contribuía para alterar completamente o sentido do que era informado, como o caso abaixo.

Figura 11.3 – Exemplo de erro de dicionário

Erro do Verificador: Tipo de erro causado por um equívoco do verificador durante a digitação. Ex: No caso abaixo, o verificador digitou um "0" inexistente, causando a divergência.

Figura 11.4 – Exemplo de erro do verificador



Erro de Grafia: Caso em que o *software* reconhecia corretamente um campo com uma palavra grafada erradamente. A divergência ocorria porque o verificador não analisava caracter a caracter, mas sim o contexto do que estava no campo, e terminava por corrigir a grafia errada.

Figura 11.5 – Exemplos de erros de grafia

GRAFIA ERRADA
Valor Proc.: TECNICA EM EMFERMAGEM
Valor Verif.: TECNICA EM ENFERMAGEM
GRAFIA RUIM
Valor Proc.: LAVOVRIA DE CAFE
Valor Verif.: LAVOURA DE CAFE

Como visto acima, o *software* reconheceu corretamente o caracter "M". É provável que o digitador nem percebesse o caracter. Desta forma, ele terminava por corrigir involuntariamente a grafia errada e provocava a divergência. Assim, o julgamento sobre a grafia era importante e mereceu atenção especial por parte dos supervisores no que diz respeito à distinção entre um erro de grafia e um erro de reconhecimento do *software*. Tal julgamento consistiu na parte mais crítica da classificação, pois a avaliação pessoal era indispensável. Tornava-se necessário separar o erro de grafia de casos em que o reconhecimento deveria ser classificado como correto.

Uso do Dicionário: Tipo de erro em que a divergência ocorria, pois, enquanto o dicionário corrigia um erro de grafia do recenseador, o verificador era orientado a não corrigir. Portanto, tratava-se de um caso particular em que nenhum dos agentes envolvidos (*software* e o verificador) poderia ser responsabilizado.

Figura 11.6 – Exemplo de erro de uso do dicionário

No caso acima, o dicionário alterou para "Romana", corretamente, e a divergência foi causada, pois o digitador seguiu o que foi escrito. Na verdade, como a recomendação passada era de que o verificador digitasse o que estivesse vendo, o verificador, e tampouco o *software* poderiam ser responsabilizados.

Já que a obtenção da taxa de divergência líquida é dependente da classificação descrita acima, foi necessário um trabalho de uniformização de critérios, realizado pelo serviço de treinamento do CCD do Rio de Janeiro, o qual contou com a presença dos Supervisores de Controle de Qualidade dos demais centros de captura, além de visitas a todos os centros para observação do andamento dos trabalhos e discussão de eventuais casos duvidosos. Em avaliações objetivas realizadas com todos os supervisores, concluiu-se que o nível de homogeneidade de conceitos estava satisfatório.

Após as etapas acima descritas, era gerada uma planilha do controle de qualidade em formato Excel, que podia ser trabalhada pelo *software* estatístico SPSS. A partir de então, eram calculadas as seguintes taxas, para cada uma das amostras selecionadas:

$$TDL = \frac{divergentes - EV\text{-}EG\text{-}UD}{reconhecidos} = \frac{ER + ED}{reconhecidos}$$

$$P_R = \frac{reconhecidos}{preenchidos}$$

$$P_{Erro\ i} = \frac{total\ do\ erro\ i}{reconhecidos}$$

Onde:

TDL: taxa de divergência líquida;

EV: erros do verificador;

EG: erros de grafia;

UD: erros no uso do dicionário;

ER: erro de reconhecimento;

ED: erro do dicionário;

P_R: proporção de reconhecidos; e

$P_{Erro\ i}$: proporção do erro do tipo *i*, *i = ER, EV, EG, UD e ED.*

As taxas eram calculadas separadamente para campos alfanuméricos e numéricos. A proporção de campos reconhecidos e a proporção dos tipos de erros eram apresentadas em gráficos de barra.

Como já mencionado, o parâmetro central do Controle de Qualidade do Reconhecimento foi a taxa de divergência líquida. Semanalmente, gráficos de controle eram feitos e analisados, para campos alfanuméricos e numéricos, em cada um dos CCDs. Considerando que o número de campos reconhecidos em cada amostra era variável, foi utilizado o gráfico de controle P (fração de não-conformidades) com limites de controle variáveis, dados por:

$$L\,(I, S) = \overline{P}\ \mu\, 3\sqrt{\frac{\overline{P}\left(1\text{-}\overline{P}\right)}{n_i}}$$

Onde:

$\overline{P}$: taxa média de divergência líquida observada na semana; e

n_i : número de campos reconhecidos na *i-ésima* amostra.

Essa fórmula implica que o comprimento dos limites de controle é inversamente proporcional à raiz quadrada do tamanho da amostra. Como visto, eram retiradas amostras para diferentes tipos de modelos de questionário. Assim, os totais de pontos amostrais grafados por dia ficaram da seguinte forma: 6 pontos para o processo de reconhecimento de campos numéricos presentes em questionários dos tipos amostra e básico e 3 pontos para o processo de reconhecimento de campos alfanuméricos, presentes apenas em questionários do tipo amostra.

Para cada amostra, uma estatística $\hat{p}$ era grafada. O $\overline{p}$ era calculado com base nos dados de uma semana. Tão logo a estatística $\hat{p}$ ficasse dentro dos limites de controle, e a seqüência de pontos não apresentasse qualquer padrão não-aleatório, inferia-se que o processo de reconhecimento apresentava controle estatístico sob o nível $\overline{p}$.

Contudo, se algum ponto amostral ficasse fora dos limites de controle, ou algum padrão não-aleatório fosse verificado, muito provavelmente, a um nível de significância de 5%, podia-se inferir que o processo de reconhecimento estivesse fora de controle estatístico, com mudança do nível de divergência para um nível mais elevado. Como as médias são próximas a 0%, trabalhou-se apenas com o limite superior de controle (LSC). Desta forma, caso o ponto amostral superasse esse limite, os supervisores investigavam a causa e a transmitiam ao Coordenador do CCD em questão.

Um procedimento complementar era identificar a natureza do erro de reconhecimento causador da elevada taxa. Tratava-se de uma segunda classificação. Na primeira, relatada acima, procurava-se dissociar o erro de reconhecimento do *software* dos demais erros, de forma a se chegar à divergência líquida. Essa segunda etapa era a tipificação do erro provocado pelo *software.* Essa tipificação era realizada apenas com imagens correspondentes aos sinais de alerta, que eram os pontos amostrais que ultrapassavam o limite de controle no gráfico.

A verificação dos tipos de erro de reconhecimento do *software,* ocorridos quando um ponto ficava além do LSC, foi iniciada no Rio de Janeiro em 23 de março de 2001. Nos demais CCDs, somente a partir da terceira semana de abril, em função da ausência nestes do *software* de consulta para a localização e classificação das imagens. Os tipos de erro de reconhecimento considerados foram os seguintes:

Erro de reconhecimento clássico: o *software* reconhecia errado um caracter sem nenhuma razão aparente. Por exemplo: vê-se normalmente um "1" e o valor processado é "5".

Erro de reconhecimento excessivo: como já mencionado, em vez de enviar o caracter para verificação, o *software* reconhecia, e errava, em função de uma imagem ou uma grafia fraca, bordas dos campos ou qualquer tipo de sujeira.

Erro do Dicionário: como já visto, esse erro ocorria quando a atuação do dicionário era suficiente para mudar um caracter ou mesmo o sentido daquilo que foi informado. Ex: "tirador de foto" reconhecido como "tirador de moto".

Omissão: tipo de erro de reconhecimento em que o *software* não reconhecia nem enviava um determinado caracter para verificação. Ex: vê-se "35" e o processado é "3", sendo o "5" simplesmente ignorado.

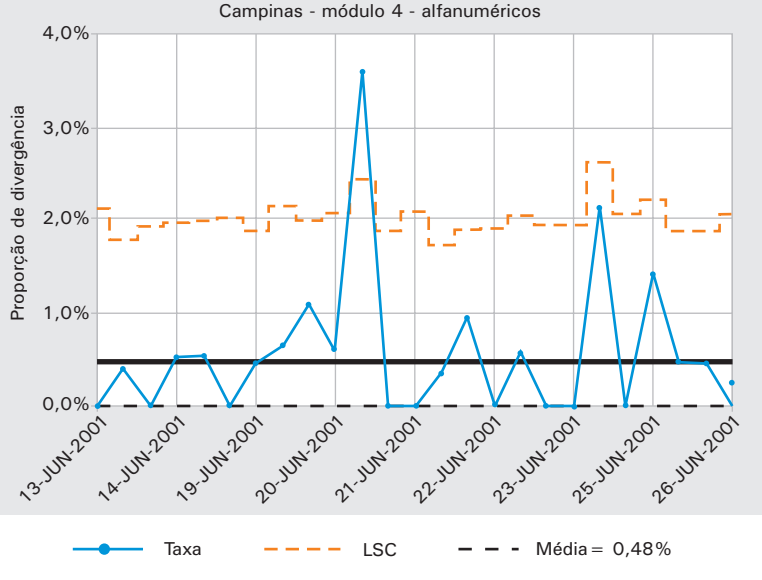
Apresenta-se, a seguir, um exemplo ilustrativo do procedimento relatado acima.

O primeiro passo seguido era a confecção do gráfico de controle. No gráfico 11.1, apresenta-se um exemplo para o módulo 4 de Campinas. Os dados são referentes aos dias 9 a 15 de maio. Como já mencionado, os limites eram variáveis, em função dos diferentes tamanhos da amostra. Observa-se que, no dia 15, aparece um ponto com uma elevada taxa (» 3,5%), que superou o limite superior de controle, significando, muito provavelmente, que o processo saiu de controle estatístico durante o período considerado.

O procedimento seguinte, acertado com os supervisores, era o da identificação do tipo de erro de reconhecimento do *software*. Para tal, os supervisores consultavam as imagens formadoras da elevada taxa e classificavam o tipo de erro de reconhecimento causador da divergência. Os resultados da investigação, no módulo 4 de Campinas, foram os seguintes:

Gráfico 11.1 - Gráfico de controle de Campinas

Campinas - módulo 4 - alfanuméricos



Quadro 11.2 - Exemplo do processo de investigação realizado pelos supervisores - 2000

Módulo	Turno	Dia	Modelo	Campo	Lote	Erros de reconhecimento			
						Clássico	Excessivo	Dicionário	Omissão
SP4	2	20/jun	Amostra	ALFA	41027177	1	0	0	0
					41027185	2	0	0	0
					41027192	1	0	0	0

Como pode ser observado, o erro de reconhecimento clássico foi o responsável pela taxa de 3,5%, aproximadamente. Além da classificação do tipo de erro, ficavam registrados o número dos lotes e o turno em que as imagens divergentes ocorriam.

Por fim, todas as informações (tabelas com a proporção de campos reconhecidos e com os tipos de erro de reconhecimento, gráficos de controle e de Pareto) eram divulgadas semanalmente na Intranet do IBGE, relatórios esses disponíveis a todos os interessados no SIGP.

Processo de seleção da amostra

As amostras em cada turno de trabalho eram formadas por 27 questionários do tipo amostra e 43 questionários do tipo básico, que eram conglomerados de campos, estes, na realidade, as unidades de observação para o controle de qualidade. O tamanho da amostra, entendido como o número de campos reconhecidos, era, portanto, variável a cada seleção realizada. Desta forma, embora os gráficos de controle P já contemplem essa restrição, torna-se importante um estudo do grau de variação do tamanho da amostra. As tabelas seguintes mostram as estatísticas básicas e o coeficiente de variação do tamanho da amostra de campos alfa e numéricos em cada um dos módulos. Para compreender a notação, a sigla da unidade da federação, seguida de um número, representa um módulo de trabalho do Centro de Captura de Dados da unidade da federação em questão.

Tabela 11.2 - Estatísticas para o tamanho de amostra de campos alfanuméricos no Controle de Qualidade do Reconhecimento - 2000

Módulo	Média	Mediana	Mínimo	Máximo	Desvio padrão	Coeficiente de variação
RJ	102,4	87,0	2,0	274,0	53,6	0,523
SP1	169,6	173,0	23,0	293,0	44,2	0,261
SP2	163,4	167,0	8,0	291,0	53,1	0,325
SP3	175,4	173,0	71,0	319,0	37,9	0,216
SP4	170,4	176,0	11,0	273,0	47,1	0,276
PR1	166,3	166,0	67,0	250,0	31,6	0,190
PR2	168,0	166,0	16,0	287,0	34,9	0,208
PR3	165,7	166,0	32,0	251,0	35,4	0,213
PR4	170,4	173,0	16,0	256,0	32,2	0,189
GO1	179,7	177,0	7,0	336,0	47,6	0,265
GO2	179,3	170,5	1,0	554,0	65,6	0,366
PB1	160,3	162,0	4,0	309,0	35,3	0,220
PB2	162,5	163,0	22,0	338,0	37,0	0,227
PB3	161,0	164,0	9,0	273,0	38,6	0,240
PB4	163,7	165,0	8,0	258,0	33,4	0,204

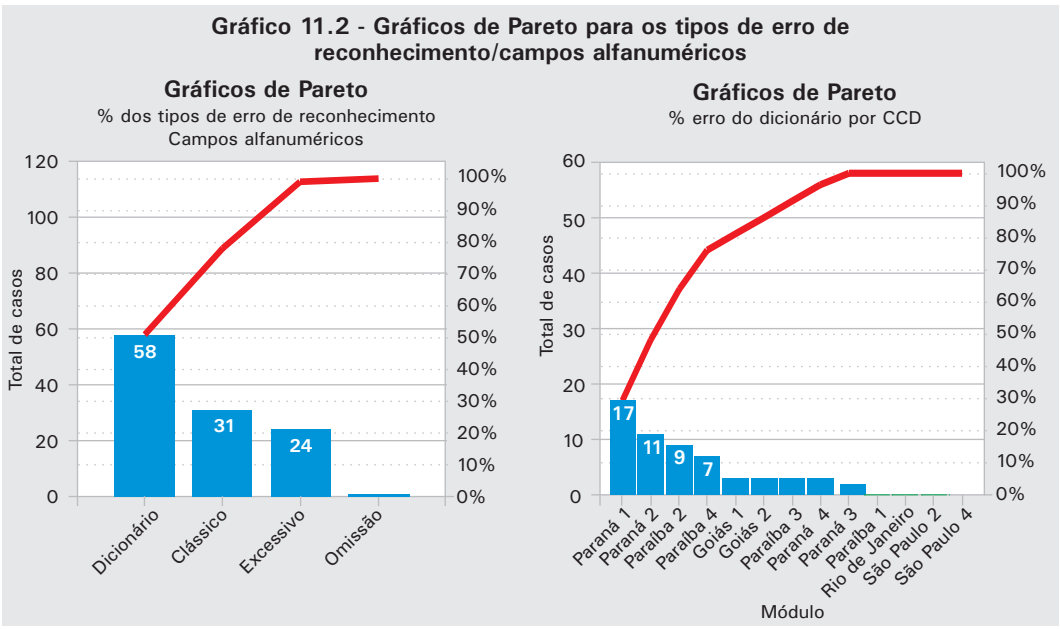
Tabela 11.3 - Estatísticas para o tamanho de amostra de campos numéricos no Controle de Qualidade do Reconhecimento - 2000

Módulo	Média	Mediana	Mínimo	Máximo	Desvio padrão	Coeficiente de variação
RJ	948,3	1 021,0	103,0	1 234,0	231,3	0,244
SP1	938,0	980,0	132,0	1 275,0	207,7	0,221
SP2	900,0	957,0	24,0	1 250,0	234,0	0,260
SP3	979,9	992,0	23,0	1 297,0	160,3	0,164
SP4	903,6	963,5	45,0	1 263,0	244,4	0,270
PR1	976,2	984,0	25,0	1 279,0	138,7	0,142
PR2	986,1	999,5	159,0	1 262,0	141,0	0,143
PR3	963,6	980,0	49,0	1 305,0	165,7	0,172
PR4	980,2	1 002,0	29,0	1 330,0	170,4	0,174
GO1	935,6	976,5	16,0	1 278,0	214,3	0,229
GO2	967,9	978,5	26,0	2 648,0	266,7	0,276
PB1	978,5	1 006,5	52,0	1 244,0	167,4	0,171
PB2	987,8	1 000,0	69,0	1 662,0	158,7	0,161
PB3	970,8	1 015,0	52,0	1 257,0	201,4	0,207
PB4	997,6	1 018,0	17,0	1 269,0	151,6	0,152

Ressalta-se apenas o elevado coeficiente de variação do número de campos alfanuméricos selecionados no Rio (cerca de 52%).

No gráfico de Pareto, representam-se os problemas relacionados à qualidade de um processo em barras verticais. A informação é colocada de forma a tornar evidente a priorização de problemas e projetos. Enquanto no eixo vertical da esquerda tem-se o total de ocorrências, no eixo vertical da direita tem-se a percentagem de cada um dos problemas encontrados. Em suma, procura-se identificar os problemas vitais, que resultam em grandes perdas para um determinado processo.

Gráfico 11.2 - Gráficos de Pareto para os tipos de erro de reconhecimento/campos alfanuméricos



Pelo gráfico da esquerda, observa-se que, em quase 60% dos sinais (pontos além do LSC), a divergência foi causada pelo erro do dicionário do *software* de reconhecimento. Pelo gráfico da direita, verifica-se que esse tipo de erro foi mais freqüente nos módulos 1 e 2 do Paraná. Cerca de 40% dos casos de erro do dicionário ocorreram nesses módulos.

11.1.6.2 Digitação ou Verificação

O controle de qualidade da verificação teve como objetivo controlar o trabalho dos digitadores que eram responsáveis por digitar os campos que não eram reconhecidos automaticamente pelo sistema de reconhecimento. Essa etapa foi denominada de Etapa de Verificação. Nela, cada campo do questionário não-reconhecido era digitado pelo menos duas vezes e, em caso de divergência entre as duas verificações, esse campo era digitado uma terceira vez, sendo esta considerada como correta.

Metodologia

Da mesma forma que no Controle de Qualidade do Reconhecimento, o Controle de Qualidade da Verificação - CQV teve como fonte de consulta o banco de dados gerados pelo sistema de reconhecimento. As variáveis utilizadas foram as seguintes:

- data: data em que foi realizada a verificação;
- verificador: identificação do verificador;
- campo: identifica o tipo de campo, alfanumérico ou numérico;
- campver: número total de campos verificados por cada verificador;
- campdiv: número total de divergências de cada verificador em relação a todos os outros;
- tcampver: número total de campos verificados em um dia, que é o somatório dos casos contabilizados em "campver"; e
- tcampdiv: número total de divergências contabilizadas em um dia. Trata-se do somatório dos casos contabilizados em "campdiv".

O Controle de Qualidade da Verificação - CQV teve, com variável de interesse, a taxa de divergência diária entre os verificadores. Um campo divergente era computado toda vez que o trabalho feito na primeira verificação não conferia com o da segunda verificação.

A taxa diária de divergência era dada pela seguinte razão:

$$TDD = \frac{tcampdiv}{tcampver}$$

Como no CQR, também foram utilizados gráficos P para a proporção de divergência diária relativa a campos alfanuméricos e numéricos. Os limites de controle também eram variáveis a cada dia, pois o tamanho da amostra (número de campos verificados no dia) também variava.

Ressalta-se que a taxa era obtida a partir do total de campos com divergência de cada verificador em relação a todos os outros. Logo, caso algum ponto ficasse além dos limites de controle, a orientação para eventuais correções foi, a princípio, repassada para o grupo de verificadores.

Foram utilizados, ainda, gráficos de dispersão, com o intuito de se ter idéia do grau de variação da taxa diária de divergência e, conseqüentemente, verificar a homogeneidade do trabalho dos verificadores durante um determina-

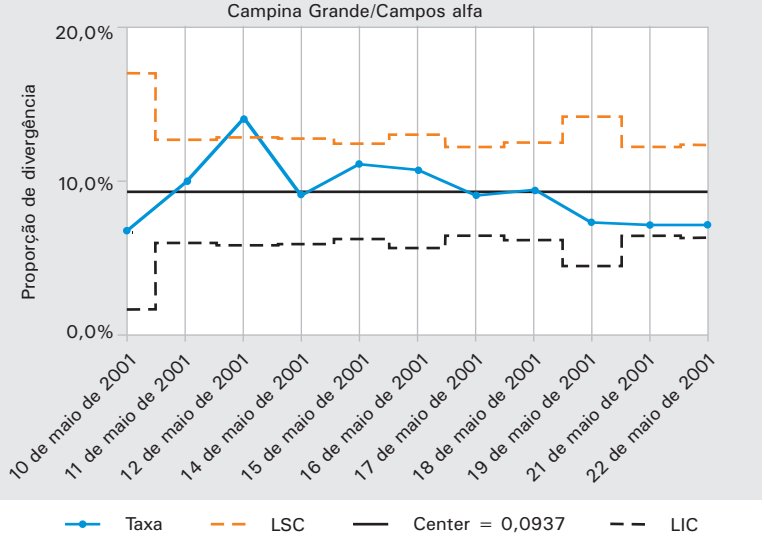
do período (uma semana, por exemplo). Para o acompanhamento da gerência e dos verificadores, os gráficos foram divulgados diariamente pelos Supervisores do Controle de Qualidade.

O procedimento utilizado é rapidamente ilustrado a seguir.

O Gráfico 11.3, apresentado em seguida, corresponde às taxas diárias de divergência dos verificadores do CCD de Campina Grande.

Gráfico 11.3 - Taxas de divergência na verificação

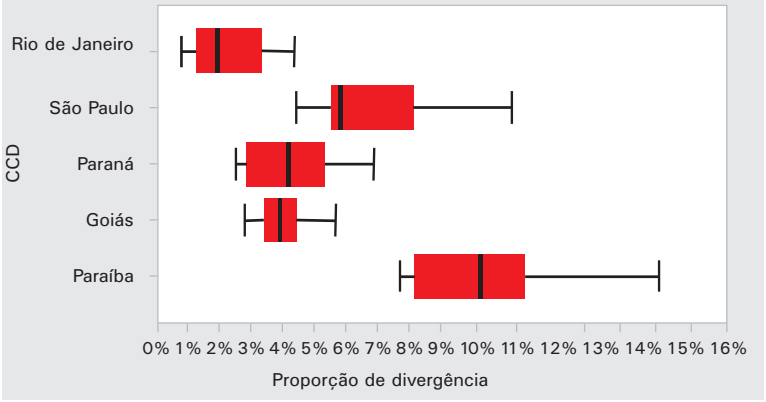
Campina Grande/Campos alfa



Como pode ser visto, o processo de verificação saiu do controle estatístico durante o período considerado, pois um ponto superou o LSC. Ao contrário do CQR, aqui já há o LIC (limite inferior de controle). Entretanto, nesse caso, um ponto abaixo do LIC nada mais significaria que uma queda do nível de divergência, o que seria um fenômeno desejável. Eventuais problemas encontrados nesses dias atípicos eram investigados e relatados. O procedimento continuava com a realização de gráficos de dispersão, conhecidos como "gráficos Box Plot". Em seguida, as taxas diárias de verificação eram representadas nos gráficos de dispersão, como o apresentado abaixo.

Gráfico 11.4 - Box Plot
Taxas diárias de divergência na 1ª verificação

Campos alfanuméricos/11-22/05



Quanto mais largo o gráfico de dispersão, significa que as taxas de divergência e o trabalho de verificação foram mais heterogêneos durante o período considerado. Conseqüentemente, quanto mais apertado o referido gráfico, mais homogêneas são as taxas e o trabalho de verificação no período. Uma função complementar do gráfico foi a de comparar o desempenho dos verificadores de diferentes CCDs. Caso fosse registrada alguma diferença entre os CCDs, a orientação natural era a investigação e o registro de eventuais causas. Todos os gráficos bem como os resultados das investigações foram apresentados semanalmente no SIGP.

Processo de seleção da amostra

Como no CQR, é importante uma rápida análise sobre o tamanho da amostra de campos alfanuméricos e numéricos. Seguem-se as tabelas com os tamanhos médios e os gráficos de dispersão dos tipos de campos selecionados.

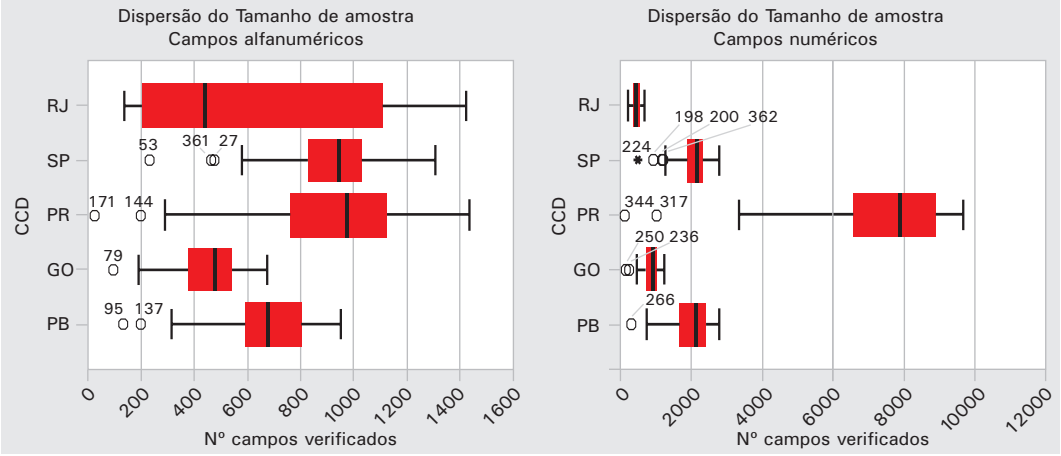
Tabela 11.4 - Estatísticas para o tamanho de amostra de campos alfanuméricos no Controle de Qualidade de Verificação - 2000

Centros de Captura de Dados	Média	Mediana	Mínimo	Máximo	Desvio padrão	Coeficiente de variação
RJ	603,8	438,5	137,0	1 421,0	459,5	0,761
SP	917,3	944,0	231,0	1 307,0	231,4	0,252
PR	910,6	972,0	25,0	1 436,0	314,2	0,345
GO	450,3	476,0	95,0	675,0	124,6	0,277
PB	662,7	678,0	132,0	951,0	184,3	0,278

Tabela 11.5 - Estatísticas para o tamanho de amostra de campos numéricos no Controle de Qualidade de Verificação - 2000

Centros de Captura de Dados	Média	Mediana	Mínimo	Máximo	Desvio padrão	Coeficiente de variação
RJ	444,1	438,5	208,0	689,0	128,7	0,290
SP	2 034,4	2 150,5	510,0	2 784,0	489,8	0,241
PR	7 362,4	7 904,0	134,0	9 678,0	2 247,8	0,305
GO	852,0	902,0	166,0	1 245,0	247,6	0,291
PB	2 003,7	2 115,0	301,0	2 795,0	577,7	0,288

Gráfico 11.5 – Box plot – dispersão do tamanho da amostra



No CQV, retirava-se uma amostra de campos alfanuméricos e numéricos trabalhados em cada um dos módulos. Como a variável de interesse era a taxa de divergência diária do CCD, esperava-se naturalmente que o tamanho de amostra variasse, dependendo do CCD. Assim, justifica-se, por exemplo, o gráfico de dispersão com os menores tamanhos de amostra em CCDs com menos módulos, como em Goiás e no Rio de Janeiro. Nos gráficos de dispersão de campos numéricos, a situação do Paraná se destaca porque dois módulos estavam configurados para selecionar mais questionários do que havia sido determinado, segundo justificativa dos supervisores locais.

Por fim, cabe ressaltar novamente a elevada variância do tamanho de amostra de campos alfanuméricos selecionados (coeficiente de variação de 76%) no CCD do Rio de Janeiro. Como este foi o primeiro a encerrar os trabalhos de apuração, os questionários do tipo amostra (que contêm os campos alfanuméricos) tornaram-se escassos, provocando a grande oscilação no tamanho de amostra.

Comentários

O Controle de Qualidade assumiu um importante papel no acompanhamento e na identificação de causas especiais presentes nos processos de reconhecimento e verificação.

No Controle de Qualidade do Reconhecimento, a confiabilidade das inferências realizadas foi assegurada com a criação de uma rotina de classificação dos erros causadores de divergência e de priorização dos lotes. Como a variável de interesse era a relação entre os casos de divergência líquida e o número de campos reconhecidos, necessitava-se que as divergências provocadas pelo *software* (erro de reconhecimento e dicionário) fossem dissociadas daquelas provocadas por outros fatores (erro do verificador, erro de grafia ou uso do dicionário) e, ao mesmo tempo, que o tamanho da amostra fosse consolidado com a passagem do último questionário no processo, fato conseguido com a priorização dos lotes.

Foram verificadas algumas distorções significativas no CQR, através dos gráficos de controle. Com a investigação da natureza do erro de reconhecimento ocorrido, foi possível identificar, por exemplo, ajustes incorretos do *scanner* para determinados lotes e a inatividade do dicionário do *software* de reconhecimento. No reconhecimento de campos alfanuméricos, o erro mais freqüente foi o de atuação do dicionário. No reconhecimento de campos numéricos, onde o gráfico acusou uma grande instabilidade estatística do processo, a divergência era causada mais freqüentemente pelo reconhecimento excessivo do *software*, ao não enviar o campo para verificação. Assim, medidas de correção do *software* de reconhecimento eram necessárias. Essas medidas demandariam eventuais ajustes que, por exigirem maior grau de sofisticação e flexibilidade, tornaram-se inviáveis.

No Controle de Qualidade da Verificação, a variável de interesse foi a taxa de divergência diária entre os verificadores. As análises estatísticas foram realizadas com a adoção de gráficos de controle e de dispersão. Realizaram-se Análises de Variância, que atestaram diferenças estatisticamente significativas entre os CCDs, nos processos de verificação de campos alfanuméricos e numéricos. Inicialmente, os CCDs da Paraíba e de São Paulo apresentaram as maiores distorções. Na Paraíba, elevadas taxas de divergência eram causadas por deficiência no critério ministrado aos verificadores, para a identificação dos caracteres. A orientação, equivocada, era analisar cada caracter, verificar casos de

grafia ruim ou errada e confirmá-los. Os supervisores locais, depois de alertarem os verificadores e repassaram o critério de verificação correto, onde cada verificador trabalharia o contexto do campo, inicialmente recomendado e seguido nos demais CCDs. Posteriormente, com o acompanhamento, as taxas estabilizaram e não foram encontradas diferenças significativas na verificação de campos alfanuméricos. No processo de verificação de campos numéricos, os gráficos de controle apresentaram, predominantemente, estabilidade estatística. Embora ocorressem diferenças significativas entre os CCDs, essas eram devidas à variação na qualidade do material trabalhado.

Por fim, considera-se oportuno apresentar um quadro geral dos processos de reconhecimento e verificação – Figura 11.7. Complementando, as tabelas 11.6 e 11.7 trazem as médias finais de divergência e de campos reconhecidos verificadas no CQR e no CQV, considerando todos os dados, desde março de 2000.

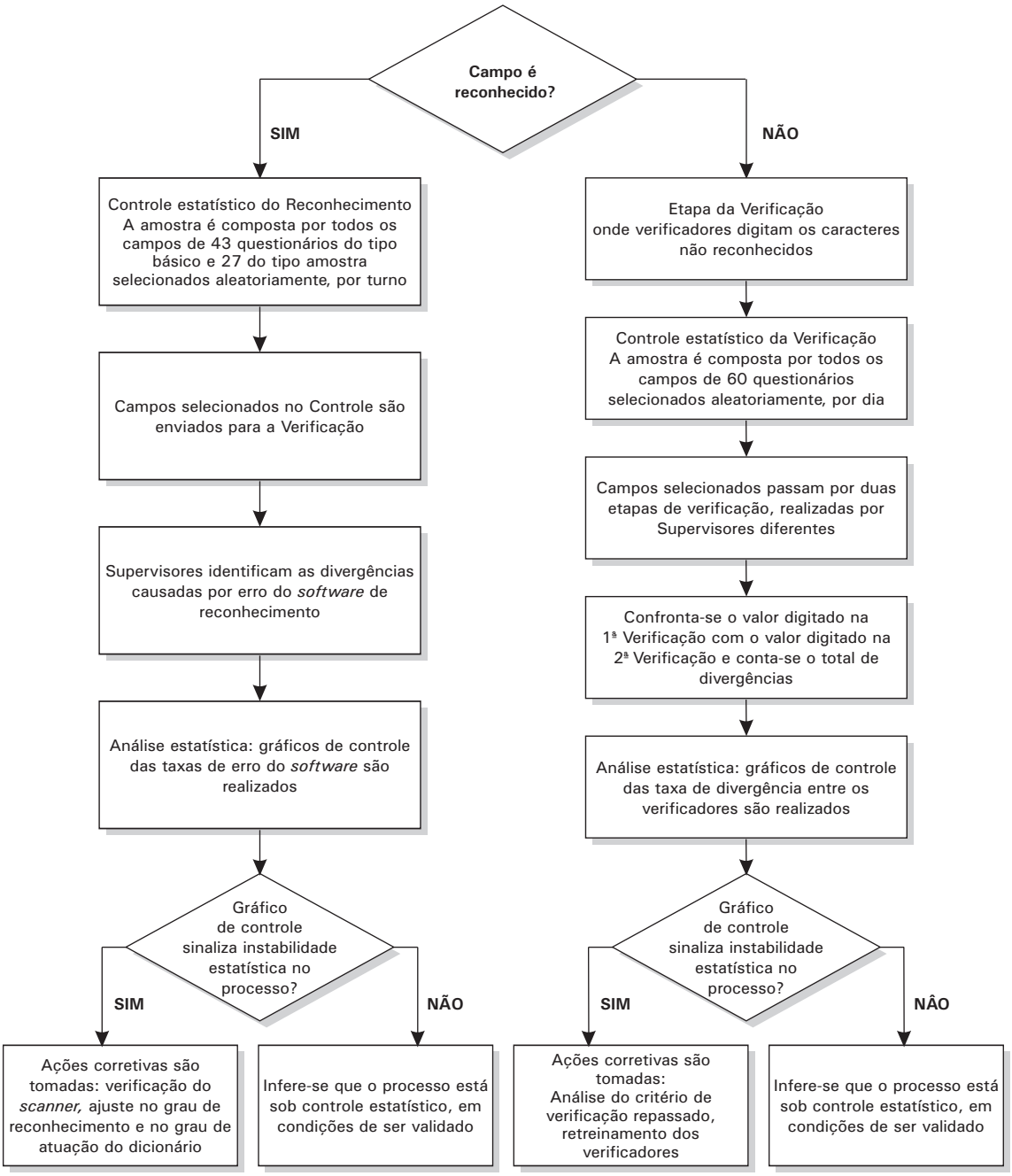
Tabela 11.6 - Proporções médias finais de divergência e de reconhecimento no Controle de Qualidade do Reconhecimento - 2000

Módulo	Proporções médias finais			
	Campos alfanuméricos		Campos numéricos	
	Percentual de divergência (%)	Percentual de reconhecidos (%)	Percentual de divergência (%)	Percentual de reconhecidos (%)
RJ	0,84	45,01	0,13	91,12
PR1	0,42	80,35	0,16	93,25
PR2	0,51	77,05	0,14	90,26
PR3	0,50	79,60	0,13	93,60
PR4	0,44	80,54	0,21	93,89
GO1	0,49	81,77	0,14	90,00
GO2	0,48	80,47	0,16	92,16
PB1	0,68	79,83	0,19	92,27
PB2	0,54	79,99	0,19	89,79
PB3	0,66	81,11	0,22	90,31
PB4	0,61	81,47	0,14	91,51
SP1	0,86	77,93	0,20	91,56
SP2	1,14	76,00	0,20	89,28
SP3	0,78	78,61	0,18	93,40
SP4	0,78	80,24	0,15	94,15
Média geral	0,65	77,37	0,17	91,76

Tabela 11.7 - Proporções médias finais de divergência no Controle de Qualidade de Verificação - 2000

Centros de Captura de Dados	Proporções médias finais de divergência (%)	
	Campos alfa	Campos numéricos
RJ	3,04	0,91
SP	4,71	0,92
PR	3,58	0,29
GO	3,96	0,70
PB	6,20	1,24
Média geral	4,44	0,82

Figura 11.7 - Controle de Qualidade do Reconhecimento/Verificação


Campo é reconhecido?
SIM
NÃO
Controle estatístico do Reconhecimento
A amostra é composta por todos os campos de 43 questionários do tipo básico e 27 do tipo amostra selecionados aleatoriamente, por turno
Campos selecionados no Controle são enviados para a Verificação
Supervisores identificam as divergências causadas por erro do *software* de reconhecimento
Análise estatística: gráficos de controle das taxas de erro do *software* são realizados
Gráfico de controle sinaliza instabilidade estatística no processo?
SIM
NÃO
Ações corretivas são tomadas: verificação do *scanner,* ajuste no grau de reconhecimento e no grau de atuação do dicionário
Infere-se que o processo está sob controle estatístico, em condições de ser validado
Etapa da Verificação
onde verificadores digitam os caracteres não reconhecidos
Controle estatístico da Verificação
A amostra é composta por todos os campos de 60 questionários selecionados aleatoriamente, por dia
Campos selecionados passam por duas etapas de verificação, realizadas por Supervisores diferentes
Confronta-se o valor digitado na 1ª Verificação com o valor digitado na 2ª Verificação e conta-se o total de divergências
Análise estatística: gráficos de controle das taxa de divergência entre os verificadores são realizados
Gráfico de controle sinaliza instabilidade estatística no processo?
SIM
NÃO
Ações corretivas são tomadas:
Análise do critério de verificação repassado, retreinamento dos verificadores
Infere-se que o processo está sob controle estatístico, em condições de ser validado

11.1.7 Crítica

O processo de crítica desenvolvido nos Centros de Captura de Dados – CCDs – após o término do Controle de Qualidade, teve como objetivo identificar e dar solução a possíveis incompatibilidades no arranjo dos questionários, bem como às divergências entre os totais de pessoas neles registrados e o número das efetivamente recenseadas, em determinada unidade domiciliar. Esse trabalho teve como base testes realizados com o material coletado no Censo Experimental, sendo o resultado da atividade conjunta de técnicos do IBGE, do Rio de Janeiro, e da empresa Politec, contratada para o trabalho de programação das rotinas dos sistemas informatizados dos CCDs.

A atividade de crítica esteve dividida em dois procedimentos, a pré-crítica e a crítica quantitativa, que deram tratamento, em separado, aos conjuntos de questionários, básico e da amostra, pertencentes a um setor. Esses procedimentos foram realizados em módulos de operação distintos nos CCDs, sendo que os erros investigados podiam decorrer tanto de equívocos na coleta dos dados, quanto de falhas no reconhecimento ótico de marcas e caracteres.

Previamente à execução da crítica, foram constituídos dois arquivos, referentes a cada um dos conjuntos de Questionários, Básico e da Amostra, de cada setor. O primeiro, proveniente da digitalização, reunia as imagens de cada questionário, e o outro congregava os seus dados, decorrentes do processo de reconhecimento ótico.

11.1.7.1 Treinamento e operação do sistema de crítica

Para o funcionamento desses módulos de operação da crítica, foi necessário o desenvolvimento de um sistema informatizado específico, bem como a capacitação dos técnicos que estariam envolvidos em sua operação. No que se refere às funções de crítica, a capacitação técnica dos operadores aconteceu através de treinamento presencial, ministrado por técnicos da DPE, em cada um dos CCDs.

O conteúdo da instrução foi distribuído entre noções conceituais do Censo, envolvendo o preenchimento da Folha de Coleta, do Questionário Básico e do Questionário da Amostra, bem como a parte prática relativa às tarefas do operador na execução das críticas. Nessa oportunidade, utilizaram-se alguns setores para simulação do trabalho, tendo como apoio as instruções do Manual da Crítica Quantitativa - CD 1.20 e do documento *Procedimentos técnicos nas atividades do CCD* (2002).

O treinamento operacional do sistema foi ministrado aos Auxiliares Censitários de Apuração – ACP, previamente treinados, na forma de correção dos erros, não só pela DPE, mas também pela Politec e pelos Coordenadores Técnicos dos CCDs. Durante o processo de crítica, os operadores dispunham de um quadro em forma de colunas com dez tipos de erros a serem corrigidos. Todo o procedimento era realizado através da visualização das imagens dos questionários, com recursos de zoom (aumento ou diminuição). Nessa etapa da crítica, o aplicativo foi estruturado de forma interativa, sendo que a chave de identificação do setor constava na imagem para cada setor criticado, com a função de facilitar o operador na tomada da decisão.

Quando o sistema de crítica acusava a ocorrência de erro, o operador observava no monitor os campos do questionário diretamente envolvidos. De acordo com a dinâmica do sistema, ficavam disponíveis comandos para aceitar o preenchimento dos campos, ou para efetuar a correção dos mesmos. O sistema permitia também a consulta à imagem do questionário com erro, assim como dos demais questionários do setor.

11.1.7.2 Execução da Pré-Crítica

Neste procedimento, o objetivo era verificar o arranjo correto dos questionários, tendo em conta a sua quantidade utilizada na unidade domiciliar entrevistada. O campo básico de investigação era o 1.12 – Questionário no Domicílio – sendo que a crítica levava em consideração a sua consistência, tendo como referência os registros feitos nos seguintes campos: 1.01 – Número do Questionário; 1.08 – Número na Folha de Coleta; e 1.09 – Número na Folha de Domicílio Coletivo.

O preenchimento correto do campo 1.12 acontecia quando:

- campo 1.12 = 1-ÚNICO – havia utilização de um único questionário na unidade entrevistada; e

- campo 1.12 = 2-TEM OUTROS ou = 3-É CONTINUAÇÃO – eram utilizados, respectivamente, o questionário principal e o(s) questionário(s) suplementar(es), quando mais de um questionário era utilizado na unidade domiciliar entrevistada.

Antes da execução das funções da Pré-Crítica, foi necessário preparar os questionários a serem criticados. Para tanto, levou-se em conta a definição de pessoa entrevistada, a saber: uma pessoa era considerada como entrevistada quando havia registro no questionário em, pelo menos, um dos seguintes campos: 4.01 a 4.05, no caso de tratar-se de Questionário Básico; e 4.01 a 4.07 para a situação de Questionário da Amostra. Um detalhamento da preparação dos questionários com esses campos pode ser encontrado nas fontes de informação citadas na lista de Referências, ao final do Capítulo.

Ainda nesse processo de preparação, havia o tratamento das situações de falta de registro para os campos 1.01, 1.12 e 4.01, ocasião em que eram efetuadas as seguintes mudanças:

- campo 1.01 – Número do Questionário – atribuição do valor igual ao mais alto desse campo encontrado no setor, mais uma unidade;

- campo 1.12 – Questionário no Domicílio – atribuição dos códigos 1, 2 ou 3, dependendo do valor da soma dos registros dos campos 1.10 – Total de Homens e 1.11 – Total de Mulheres;

- campo 4.01 – Sexo – atribuição do código 0, sempre que fosse atendida a definição de pessoa entrevistada e houvesse omissão da informação.

Após esse preparo, iniciava-se a Pré-Crítica que, em linhas gerais, buscava identificar os questionários únicos ou os conjuntos de questionários principal e suplementar(es) em cada uma das unidades do setor. Desta forma, eram analisadas as seguintes situações de erro:

- questionários classificados como principais, sem seu(s) correspondente(s) suplementar(es); e

- questionários suplementares, sem seus respectivos questionários principais.

Nesses casos, o operador observava na tela a relação distinta de questionários que estavam em situação de erro, os quais se distribuíam entre os que indicavam ter suplementar e aqueles que eram suplementares. Assim, recorrendo às imagens dos questionários e observando os quesitos envolvidos na crítica, era possível executar uma opção do sistema, implementando a solução desejada para um dos campos 1.12, 1.08 e 1.09.

11.1.7.3 Execução da Crítica Quantitativa

Terminada a Pré-Crítica, começava a operar o módulo da Crítica Quantitativa que, basicamente, detectava e dava solução às inconsistências existentes nos totais de questionários do setor e no total de pessoas entrevistadas na unidade domiciliar.

Ao todo, o sistema investigava a ocorrência de dez tipos de erro, e funcionava através da execução de ciclos ou passagens de crítica. Assim, uma vez submetido o conjunto de questionários ao sistema, todos os que apresentassem um determinado erro eram identificados para investigação. Em seguida, o operador passava à realização dos acertos, até que todos os questionários com erro fossem tratados. Depois disso, novamente todo o lote era submetido ao sistema de críticas, para que o próximo erro fosse objeto de investigação. Isto se repetia até que todos os dez erros fossem trabalhados e resolvidos. Na Crítica Quantitativa, foram investigadas na as seguintes situações de erro:

1 - divergência entre os registros das chaves – campos 1.02 a 1.06 - de identificação do questionário e do setor;

2 - presença de questionário com campo 1.12 igual a 2 e ausência do correspondente questionário suplementar; existência de seis pessoas entrevistadas, independentemente dos valores registrados nos campos 1.10 – Total de Homens e 1.11 – Total de Mulheres;

3 - presença de questionário com campo 1.12 igual a 2 e ausência do correspondente questionário suplementar; existência de menos de seis pessoas entrevistadas, independentemente dos valores registrados nos campos 1.10 e 1.11;

4 - existência de questionário com campo 1.12 igual a 3 e falta do correspondente questionário principal;

5 - registros nos campos 1.08 – Número na Folha de Coleta – ou nos campos 1.09 - Número na Folha de Domicílio Coletivo – iguais em mais de um questionário, que apresentavam registros diferentes para o campo 1.01 – Número do Questionário;

6 - questionário(s) em duplicata;

7 - questionário(s) em falta ou em excesso;

8 - pessoa nascida após a data de referência do Censo;

9 - total de pessoas entrevistadas no questionário, maior que a soma dos registros dos campos 1.10 e 1.11;

10 - total de pessoas entrevistadas no questionário, menor que a soma dos registros dos campos 1.10 e 1.11.

Os pontos listados a seguir procuram esclarecer melhor as situações de erro apresentadas. Assim, temos:

- na investigação do erro 1, buscava-se saber se o questionário pertencia ou não ao setor objeto de crítica; a chave de identificação correta do setor, composta pelos campos 1.02 a 1.06 – respectivamente, UF, Município, Distrito, Subdistrito e Setor – era visualizada, juntamente com a imagem do questionário; nos casos em que o operador indicasse que o questionário não pertencia àquele setor, o sistema automaticamente o excluía, arquivando-o em um banco de questionários disponíveis para, mais tarde, ter a possibilidade de ser realocado em outro setor;

- as funções de crítica relativas aos erros 2, 3 e 4 foram mantidas na Crítica Quantitativa, tendo em vista a possibilidade de equívocos do operador no procedimento de Pré-Crítica;

- o erro 5 alertava para a possibilidade de uso de mais de um questionário numa mesma unidade entrevistada, não apresentando registros iguais para os campos 1,01;

- no caso do erro 6, era importante a definição de questionário em duplicata, ou seja, dois ou mais questionários eram considerados duplicatas quando possuíssem o mesmo valor para os campos 1,01 e apresentassem situação de inconsistência nos códigos dos seus campos 1.12;

- a situação do erro 7 abrangia duas definições: havia falta de questionário(s) no setor quando ocorressem saltos na seqüência numérica dos valores dos seus campos 1.01; um questionário era considerado em excesso no setor quando o valor do seu campo 1,01 superasse o número de unidades entrevistadas no setor, registrado no SIGP;

- no erro 8 eram identificadas, preliminarmente, as pessoas nascidas após 31 de julho de 2000; a pessoa era mantida ou retirada do questionário, após ser feita a verificação da imagem correspondente ao mês e ano de nascimento – campo 4.03 no Questionário Básico e 4.05 no Questionário da Amostra – tendo em vista eliminar possíveis equívocos do reconhecimento ótico;

- as ocorrências dos erros 9 e 10 alertavam para inconsistências entre o total dos valores registrados nos campos 1.10 e 1.11 e o número de pessoas cujos registros atendiam à definição de pessoa entrevistada.

Após a execução da Crítica Quantitativa, resultava que questionários e pessoas fossem retirados ou incluídos no setor. Em função disso, o sistema emitia um relatório de análise que, entre outros propósitos, possibilitava avaliar o desempenho dos operadores da crítica, levando em conta parâmetros de controle. Desta forma, o sistema apresentava os seguintes dados:

- variação no número de unidades entrevistadas, levando em consideração o seu total inicial registrado no SIGP;

- variação no total de homens, de mulheres e no total de homens mais mulheres, em relação aos registros do SIGC;

- variação na freqüência das alterações do campo 1.12, correspondentes às mudanças dos códigos 2 e 3 para código 1, em relação ao total de entrevistas registradas no SIGP.

Observe-se que parâmetros de controle estabeleciam, em relação a essas variações, valores máximos admissíveis de 1,5%, 3% e 2% para, respectivamente, o número de unidades entrevistadas, o número de pessoas - total e por sexo – e a quantidade de alterações no campo 1.12. Caso, em um determinado setor, algum desses limites fosse superado, o Coordenador Técnico do CCD, após as investigações necessárias, aceitava o trabalho do operador ou, caso contrário, determinava que o setor fosse novamente submetido ao procedimento de crítica.

11.1.8 Transmissão dos dados e gravação final dos dados e imagens

Os setores criticados e fechados eram enviados para o computador central do IBGE no Rio de Janeiro *(mainframe*), onde era feita a validação da recepção dos dados; caso se registrasse algum tipo de erro, o CCD era automaticamente informado para providenciar a correção e retransmitir o referido setor.

Os CCDs tinham como funções fazer *backups* diários em fita, não apenas dos setores transmitidos e validados, mas também do banco de dados e de imagens de todos os setores, além de providenciar a gravação dos CDs, com as imagens dos respectivos questionários, para os setores já transmitidos e validados, a fim de que fosse possível consulta através de *software*.

11.2 Dados da Folha de Coleta

Para a captura dos dados das Folhas de Coleta, foi projetada e desenvolvida uma solução que se fundamenta na tecnologia de reconhecimento ótico de caracteres, utilizada na captura dos dados dos questionários. As Folhas de Coleta do Rio de Janeiro e da Região Norte foram apuradas no Centro de Captura de Dados do Rio de Janeiro, sendo as demais no CCD-PR. As informações das Folhas de Coleta foram armazenadas numa base de dados central, e as imagens, agrupadas por unidade da federação, foram gravadas em base de dados e CDs, com recursos para a visualização das mesmas.

11.2.1 Recepção, preparo do material e digitalização

Nesta etapa, as Folhas de Coleta foram recepcionadas, conferidas e preparadas para a digitalização, sendo gerada a capa de lote. No caso em que algum formulário estivesse num estado que impedisse a sua digitalização, o mesmo foi "passado a limpo", tendo sido seus dados transcritos para um novo formulário. Esse material apresentava condições precárias, tais como: acúmulo poeira, rasgões, manchas e o de alguns setores com contaminação de cupim. Assim, devido ao seu estado de má conservação, foi necessário tratamento especial em todas as etapas do processamento.

As Folhas de Coleta eram conferidas por identificação do setor e numeração de páginas; as folhas sem numeração ficavam mantidas sem o número, conforme orientação da DPE.

Os lotes foram criados por setor censitário. Vários setores podiam estar associados a uma única pasta física. Cada pasta recebeu uma identificação única, permitindo, assim, a fácil localização das Folhas de Coleta de qualquer setor. Posteriormente, os formulários foram digitalizados.

Nessa etapa, era verificado o preenchimento dos campos UF, Município, Distrito, Subdistrito e Setor. Se estivessem em branco, eram preenchidos com a mesma identificação do setor em que se estava trabalhando, ou seja, verificando-se a folha anterior/posterior juntamente com os campos Número da Página, Localidade e Logradouro, assim como o número de ordem dos campos Unidade Visitada, Domicílio Particular e Domicílio Coletivo. Ainda nessa conferência, também se observava se todos os campos principais, bem como os de Quarteirão ou Quadra, Face e CEP estavam preenchidos corretamente de acordo com as orientações técnicas.

As Folhas de Coleta eram contadas por setor. Verificava-se o estado desse material, se possuía sujeiras, manchas, rasgões. O controle era feito de forma rigorosa, evitando-se folhas com o mesmo número ou o preenchimento incorreto da identificação do setor. Tais folhas eram abertas e limpas com esponja para a retirada de poeira e resíduos. Já o preenchimento ilegível era controlado por operadores sob orientação técnica.

Com o material devidamente preparado, as folhas eram organizadas por setor, e vários setores eram postos em caixas. Essas caixas seguiam para serem registradas no sistema e gerar a capa de lote para cada setor. O número dado às caixas com os respctivos setores era gerado em função de um banco de dados existente nos CCDs.

As caixas com os setores eram levadas para a digitalização, acompanhadas das capas de lote; as pastas (numeradas) eram encaminhadas para a digitalização, onde o operador, através da tela do computador, comandava o início do processo.

O mesmo procedimento adotado na captura de dados dos questionários também foi utilizado para a digitalização das Folhas de Coleta. A partir da digitalização, era formado o arquivo de imagens do lote, que era enviado para o servidor de imagens.

11.2.2 Reconhecimento ótico

Neste processo, as imagens das Folhas de Coleta digitalizadas foram identificadas, e os dados relativos a marcas e caracteres foram reconhecidos através de procedimentos e *softwares* específicos. Para auxílio do reconhecimento, foram utilizadas tabelas de CEP, logradouros e setores.

Os procedimentos para reconhecimento dos campos com códigos de barras, marcas, números e alfabéticos foram semelhantes aos adotados para a captura de dados dos questionários, sendo utilizados os mesmos *softwares*, com as devidas adaptações.

11.2.3 Digitação dos campos com baixo nível de reconhecimento

Esta etapa visava identificar os formulários não-reconhecidos e também corrigir os campos que não alcançaram o índice mínimo de reconhecimento preestabelecido. Exigiu a intervenção de digitadores, para identificar os formulários, digitar, corrigir ou completar esses campos.

11.2.4 Controle de qualidade

O Controle Estatístico de Processos também foi adotado para verificar e validar a qualidade dos processos de reconhecimento e verificação das informações das Folhas de Coleta do Censo Demográfico de 2000. O presente item descreve o plano de Controle de Qualidade utilizado nesse processo. Para tanto, estão relatados o sistema de planejamento amostral e os procedimentos válidos para garantir a confiabilidade das informações obtidas. Por fim, apresentam-se as ferramentas estatísticas necessárias à análise dos resultados encontrados.

Planejamento das amostras

No planejamento amostral do processo de controle para as folhas de coleta, foram consideradas as seguintes situações:

- os objetos do Controle de Qualidade são os processos de reconhecimento e verificação das Folhas de Coleta (CD 1.07). O Controle de Qualidade do Reconhecimento (CQR) e O Controle de Qualidade da Verificação (CQV) são independentes, sendo os dados apresentados em diferentes consultas;
- a unidade de seleção é o formulário da Folha de Coleta, contudo, as unidades amostrais do CQR e do CQV são, respectivamente, campos reconhecidos e campos verificados. Conseqüentemente, o tamanho final das amostras para os campos é variável;
- foram extraídas duas amostras por dia, sendo uma para cada turno de trabalho. Cada amostra era composta por 70 formulários, selecionados a cada dois minutos. Os horários de seleção foram entre 9:00h e 11:20h (primeira amostra) e entre 13:00h e 15:20h (segunda amostra);
- no Controle de Qualidade do Reconhecimento, o domínio de estimação foi turno/módulo. Isto significa dizer que o controle de qualidade atuou analisando separadamente os dados de cada módulo duas vezes por dia (turno da manhã e turno da tarde). No Controle de Qualidade da Verificação, o domínio foi dia/CCD, ou seja, a análise foi diária para cada um dos Centros. Para cada amostra, foram registradas, em um banco de dados, as seguintes informações: módulo, data, total de folhas selecionadas, tipo de campo e turno;
- no Controle de Qualidade do Reconhecimento, além das variáveis de identificação, acima mencionadas, foram ainda registrados para cada amostra: número de campos preenchidos (*preench*), número total de campos reconhecidos (*reconhec*), número total de campos com divergência (*diverg*), número de campos reconhecidos e número de campos com divergência para cada um dos tipos de campo (UF, Localidade, Número do Logradouro, etc) e, por fim, o total de cada um dos tipos de erro causadores de divergência (erro de reconhecimento, erro do verificador e erro de grafia);
- no Controle de Qualidade de Verificação, além das variáveis de identificação mencionadas, foram registradas as seguintes variáveis: operador da primeira verificação, total de campos verificados e total de campos com divergência.

11.2.4.1 Reconhecimento ótico

Como no Controle de Qualidade do Reconhecimento, o objetivo era avaliar o desempenho do programa de reconhecimento, e foi necessário identificar apenas as divergências causadas por erro de reconhecimento do programa. Sendo assim, houve uma etapa de classificação dos tipos de erro causadores de divergência. No caso das Folhas de Coleta, as divergências poderiam ser causadas por erro de reconhecimento do programa (ER), por erro de digitação do verificador (EV) ou por erro de grafia (EG). As imagens dos campos divergentes eram disponibilizadas num aplicativo computacional para classificação dos diversos tipos de erros, a fim de contabilizar aqueles que efetivamente representavam erros de reconhecimento por parte do programa. Essa classificação era feita pelos Supervisores de Controle de Qualidade de cada Centro de Captura de Dados. Para tanto, os supervisores receberam treinamento no CCD do Rio de Janeiro no mês de abril de 2001. Além dos procedimentos repassados para

a classificação dos erros, o treinamento também consistiu na apresentação de noções básicas sobre ferramentas de controle estatístico e sobre o programa estatístico usado na análise dos resultados.

Como o objetivo era identificar apenas os casos de erro de reconhecimento, o CQR tinha como variável-chave a taxa de divergência líquida, que conseqüentemente não considerava os casos de divergência causados por grafia ruim ou erro do verificador. Essa taxa foi estimada pela razão:

$$TDL = \frac{diverg - EV - EG}{reconhec}$$

Onde:

TDL: taxa de divergência líquida;

diverg: número total de divergências;

EV: erros do verificador; e

EG: erros de grafia.

A classificação dos erros era realizada somente quando o último formulário selecionado para a respectiva amostra passasse pela etapa da verificação.

Mecanismos de análise e de ação corretiva

No controle de qualidade do reconhecimento, os gráficos de controle foram feitos e analisados semanalmente, considerando-se a taxa de divergência líquida para campos alfanuméricos e campos numéricos. Como o número de campos em cada amostra era variável, foi adotado o gráfico $\overline{P}$ com limites de controle variáveis. Os limites foram dados por:

$$L\,(I, S) = \overline{P} \pm 3\sqrt{\frac{\overline{P}\left(1\text{-}\overline{P}\right)}{n_i}}$$

onde $\overline{P}$ é a taxa média de divergência líquida observada na semana, e n_i o número de campos reconhecidos na *i-ésima* amostra.

Por dia, eram esperados três pontos para a análise do processo de reconhecimento de campos alfanuméricos e outros três pontos para a análise do processo de reconhecimento de campos numéricos. O $\overline{P}$ do gráfico era estimado com base nos dados de uma semana. Para cada amostra, uma estatística $\hat{p}$ era anotada no gráfico de controle.

Deve-se considerar o processo de reconhecimento como sob controle estatatístico, quando o gráfico de controle apresenta todos os pontos amostrais dentro dos limites de controle, além de não caracterizar qualquer padrão não-aleatório. Contudo, no caso do gráfico de controle emitir um sinal, ou seja, pontos amostrais além dos limites de controle ou que fornecem padrões não-aleatórios, o procedimento a ser seguido era a investigação de possíveis problemas no processo de reconhecimento.

Entre os procedimentos de investigação usados, estava a caracterização do tipo de erro de reconhecimento, realizada pelos Supervisores de Controle de Qualidade. Tal caracterização somente era realizada quando as imagens correspondentes ao sinal emitido pelo gráfico pudessem ser visualizadas. Identificado o problema ou a causa especial, procedia-se à possível ação corretiva. Esta deveria ser tomada, em conjunto, pela gerência do processo e o pessoal encarregado do siste-

ma computacional. Entre outras ferramentas de subsídio ao Controle de Qualidade do Reconhecimento está a confecção de gráficos de dispersão das taxas de divergência, bem como de gráficos de barra com a proporção de campos reconhecidos, acompanhados da proporção de cada um dos tipos de erro, além dos gráficos de Pareto dos erros de reconhecimento. Os relatórios eram semanalmente disponibilizados para o acompanhamento gerencial interno do processo.

11.2.4.2 Digitação ou Verificação

No Controle de Qualidade da Verificação, não houve necessidade de qualquer classificação prévia para a realização das análises. Todas as variáveis necessárias à análise já eram fornecidas diretamente pelo sistema de consulta.

Além disso, os gráficos de controle também foram o principal subsídio à análise, e a variável de interesse era a taxa de divergência diária entre os verificadores, dada pela seguinte razão:

$$TDD = \frac{\sum_i d_i}{\sum_i v_i}$$

Onde:

TDD: taxa de divergência diária;

d_i: número de divergências do verificador i; e

v_i: número de campos verificados pelo verificador i.

Como o número de campos verificados a cada dia era variável, trabalhava-se com o gráfico P com limites de controle variáveis. A interpretação é similar à mencionada no Controle de Qualidade do Reconhecimento: caso algum ponto amostral ultrapassasse os limites de controle, ou fosse identificada uma configuração de algum padrão não- aleatório, isto significaria uma indicação de processo fora de controle estatístico e, caso fosse necessário, seriam aplicadas eventuais ações corretivas. Como a taxa era obtida a partir do total de campos com divergência de cada verificador em relação a todos os outros, não havia como identificar inicialmente os causadores da divergência. Desta forma, as ações corretivas eram aplicadas sobre o grupo completo de verificadores.

Além do gráfico de controle, utilizaram-se outras ferramentas, tais como os gráficos de dispersão, importantes para verificar a homogeneidade do trabalho dos verificadores durante um determinado período, por exemplo. Os relatórios do controle de qualidade da verificação eram disponibilizados semanalmente para o acompanhamento gerencial interno do processo.

11.2.4.3 Comentários

O controle de qualidade teve um papel importante no acompanhamento, na detecção e na correção de problemas nos processos de reconhecimento e verificação das Folhas de Coleta.

No CQR, a confiabilidade das inferências era assegurada com a rotina de identificação dos agentes causadores de divergência. Essa tarefa, realizada pelos supervisores de CQ, foi necessária para se chegar a uma taxa de erro de reconhecimento confiável, desvinculando da análise as divergências provocadas por um erro do verificador ou em função de uma grafia errada.

As principais distorções verificadas foram encontradas no CQR de campos alfanuméricos. Por meio dos gráficos de controle e de dispersão, pôde-se inferir a falta de controle estatístico nesse processo. Além da instabilidade, as taxas de erro de reconhecimento chegaram a atingir níveis inaceitáveis em praticamente todos os módulos, ficando em torno de 6% no módulo do Rio de Janeiro e 9% no módulo 1 de Curitiba. Como ação corretiva, sugeriu-se a redução do grau de reconhecimento do *software*. Tais ajustes necessários foram providenciados e implementados inicialmente no Rio de Janeiro. Os ajustes apresentaram resultados satisfatórios, com a taxa de erro de reconhecimento estabilizando-se em torno de 2%.

Quanto ao CQV, a análise dos gráficos de controle e de dispersão não indicou qualquer problema específico com o trabalho de verificação. As taxas de divergência variaram em torno de 2 a 4%. Esses níveis foram considerados satisfatórios e muito próximos aos observados na etapa de verificação dos questionários do censo. Detectou-se apenas uma flutuação das taxas de divergência em função da alteração brusca na qualidade do material trabalhado. Essa característica ocasionou índices de autocorrelação em torno de 30%, capazes de comprometer os pressupostos para a utilização do gráfico P (proporções de não-conformidades). Isto ocorre porque, com grande autocorrelação, os limites de controle do gráfico P ficam subestimados e não há sentido em utilizá-lo como ferramenta de controle. Adotou-se, então, o gráfico de observações individuais para as proporções de divergência. Esse gráfico estima corretamente a largura dos limites a partir da amplitude móvel entre as proporções de divergência observadas. Desta forma, o gráfico de observações individuais tornou-se um mecanismo eficiente para o controle estatístico e a detecção de eventuais causas especiais nesse processo.

11.2.5 Crítica

Concluída a etapa de controle de qualidade, tinha início a crítica das informações da Folha de Coleta. As funções de crítica, dependendo do(s) campo(s) em trabalho, podiam dar tratamento tanto à falta de informações quanto às situações de inconsistência e de falso reconhecimento ótico, verificando-se também, em algumas situações, a compatibilidade dos registros.

A exemplo do que ocorreu na crítica dos questionários, o sistema de crítica investigava, separadamente, um conjunto de Folhas de Coleta de um setor. Todo o processo de trabalho teve o suporte de um sistema informatizado que acessava, além dos dados e das funções de crítica, as imagens das folhas, resultado da digitalização. Assim, através de um mecanismo de interação, baseado nas respostas a um conjunto de indagações do sistema, o operador ou o próprio sistema dava solução às situações de erro identificadas.

Após execução do sistema, eram apresentadas, para investigação do operador, as folhas identificadas segundo um conjunto de nove erros. Com base nisto, iniciava-se o trabalho de investigação de todos os erros do setor, em uma única passagem de crítica, diferentemente do que se fazia na crítica dos questionários, onde era investigado um único erro em cada ciclo de crítica.

Os erros 1 e 2 investigavam o campo 1.09 – Número da Página – buscando identificar equívocos na seqüência de seus valores, tendo em vista apontar registros, respectivamente, que superassem o total de páginas do setor ou que estivessem em duplicidade.

No erro 3, eram objeto de crítica os campos de identificação 1.01 a 1.05, respectivamente UF, Município, Distrito, Subdistrito e Setor. Com o auxílio da imagem, era possível conhecer se a folha fazia parte do setor e, caso o equívoco fosse decorrente de um falso reconhecimento ótico, haveria a correção do(s) valor(es). No entanto, se a folha não pertencesse ao setor examinado, encerrava-se a crítica, devolvendo o setor para exame dos técnicos da etapa de recepção e preparo do material.

O erro 4 verificava a falta de informação nos campos 1.06 – Quadra – e 1.07 – Face. Caso o lote de crítica correspondesse a um setor rural, o sistema imputava 00 – dois zeros – a esses campos; no caso de área urbana, nenhuma correção era implementada.

Os registros dos campos 1.10 – Localidade – e 1.11 – Logradouro - foram objeto de comparação com suas respectivas imagens. Caso o operador identificasse divergências, ficava caracterizado o erro 5, cuja correção acontecia com a digitação do(s) campo(s), onde devia ser respeitado, fielmente, o(s) conteúdo(s) apresentado(s) na imagem.

O campo 1.08 – CEP – foi criticado, comparando-se sua informação com a do cadastro dos Correios e, dependendo do setor, também com as do CADLOG. O erro 6 dava tratamento às situações de falta de registro ou de registro incorreto nesse campo e, quando não fosse possível encontrar um código para um determinado logradouro, fazia-se o registro por meio do CEP geral do município.

O erro 7 dava início à crítica das linhas da Folha de Coleta, e ocorria quando o campo 2.01 – Número do Logradouro – estivesse sem informação. Como nem todos os campos da linha foram objeto de digitalização, era necessário que o operador, através da observação da imagem, verificasse se toda a linha estava em branco e, em caso afirmativo, fizesse a sua exclusão.

No erro 8, o campo 2.03 – Espécie – era criticado quando estivesse em branco. Neste caso, o sistema indagava ao operador se existiam preenchimentos válidos para os campos 2.08, 2.09 e 2.10, que registravam, respectivamente, o nome do responsável, o número de questionários preenchidos e o de pessoas recenseadas. Se a resposta do operador fosse positiva, e dependendo das informações dos campos 2.05 e 2.06 - respectivamente, Número de ordem – Domicílio Particular e Número de Ordem – Domicílio Coletivo – era possível recuperar para o campo 2.03 os códigos 1 – Domicílio Particular Ocupado -, 5 – Domicílio Coletivo –,ou 6 – Não-Residencial.

Por fim, o erro 9 tratava as situações dos campos 2.05 e 2.06 quando, simultaneamente, apresentavam, ou não, preenchimento. Cabe registrar que quando houvesse valor para o campo 2.05, não deveria existir para o 2.06 e vice-versa. Durante a investigação desse erro, dependendo da situação tratada, poderia ser revisto, também, o registro do campo 2.03.

11.2.6 Transmissão e gravação dos dados e imagens

Os dados das Folhas de Coleta, obtidos no processo de captura, foram gravados em banco de dados. Para cada setor (lote de trabalho), foram formatados, preparados e transmitidos para o equipamento central do IBGE, no Rio de Janeiro. De acordo com as rotinas de segurança, relativamente aos setores processados, eram feitas cópias dos dados em fitas e CDs.

As imagens dos setores validados eram transferidas para o servidor central de imagens do Centro de Captura de Dados. As imagens gravadas, agrupadas por unidade da federação, com a indexação dos setores correspondentes, foram gravadas em CDs, com recursos para a visualização das mesmas.

Referências

ALBIERI, S.; DIAS, A. J. R.; MENEZES, A. C. F.; GREEN, A. P. L. *Controle de qualidade da captura de dados do censo 2000*: especificações para o planejamento das amostras: 2ª versão. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2001.

BRAVO, P. C. Elementos de controle estatístico de qualidade. In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 6., 1984, Rio de Janeiro. *Atas...* Rio de Janeiro: UFRJ, 1984. 585 p.

CENSO demográfico 2000: manual da crítica quantitativa – CD 1.20. Rio de Janeiro: IBGE, 2000.

COCHRAN, W. G. *Sampling techniques*. 3rd ed. New York: Wiley, c1977. 428 p.

GUIMARÃES, N. R. *Controle de qualidade do censo 2000:* análise final dos processos de reconhecimento e verificação. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2001. 54 p.

_____. *Descrição do plano de controle de qualidade do reconhecimento e da verificação do processo de captura dos dados das folhas de coleta*. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2002. 24 p.

HRADESKY, J. L. *Productivity and quality improvement:* a practical guide to implementing statistical process control. New York: McGraw-Hill, 1988. 243 p.

MONTGOMERY, D. C. *Introduction to statistical quality control.* 3rd ed. New York: Wiley, 1997. 742 p.

_____. *Introduction to statistical quality control*. 4th ed. New York: Wiley, 2001. 796 p.

PROCEDIMENTOS técnicos nas atividades do CCD. IBGE, Diretoria de Informática, 2002. 10 p.

WERKEMA, M. C. C. *Ferramentas estatísticas básicas para o gerenciamento de processos*. Belo Horizonte: UFMG, Escola de Engenharia: Fundação Christiano Ottoni, 1995. 384 p.

_____. *Avaliação da qualidade de medidas*. Belo Horizonte: UFMG, Escola de Engenharia: Fundação Christiano Ottoni, 1996. 101 p. (Série ferramentas da qualidade, 13).

12 Apuração dos dados

Uma vez encerrada a etapa de captura de dados, com sua armazenagem em meio magnético, teve início a etapa de apuração que, em síntese, destinava-se ao tratamento de todas as informações coletadas, preparando-as para a divulgação e disseminação.

Este capítulo descreve todas as etapas de trabalho, começando pelo processo de apuração dos Resultados Preliminares e da Sinopse Preliminar, que tiveram como fonte os dados das Cadernetas dos Setores, chegando à apuração dos Questionários Básicos e da Amostra que deram origem às divulgações dos Resultados do Conjunto Universo e Resultados da Amostra. Contém os procedimentos de: aplicação e verificação de códigos; de crítica, imputação e expansão dos dados, validação dos resultados e tabulação.

Cada uma destas etapas encerra um grau de complexidade elevado. Assim, constituem-se como fatores condicionantes à criação de mecanismos de controle e avaliação de cada uma delas, bem como um planejamento adequado que permita seu encadeamento de forma sincronizada, a fim de garantir o cumprimento de prazos e um padrão de qualidade adequado do processo de apuração.

12.1 Resultados Preliminares

Em dezembro de 2000, os primeiros resultados foram apresentados ao público com a publicação *Censo Demográfico 2000: resultados preliminares*. Divulgada após duas semanas

de encerrada a coleta dos dados, a publicação foi o primeiro retrato da população brasileira e teve, além de outros objetivos, fornecer subsídios ao Tribunal de Contas da União para o estabelecimento das cotas do Fundo de Participação dos Estados e do Fundo de Participação dos Municípios.

O Sistema de Indicadores Gerenciais da Coleta - SIGC foi a fonte de dados para a publicação. As informações correspondentes ao resumo do CD 1.06 – Caderneta do Setor foram disponibilizadas no sistema em diferentes datas, à medida que cada unidade da federação encerrava a sua coleta. Em função disso, os resultados tiveram como referência o dia 11 de dezembro de 2000, data da última atualização no sistema, tendo caráter preliminar, diferindo dos resultados definitivos divulgados posteriormente pelo IBGE.

A publicação *Censo Demográfico 2000: resultados preliminares* constou de uma parte impressa e outra em CD-ROM. O volume impresso continha os comentários sobre a evolução do crescimento da população, mapas da densidade populacional do Brasil, dos estados e de seus municípios, além de tabelas da população recenseada, para todos os municípios do País, segundo o sexo e a situação do domicílio. No CD-ROM, além das tabelas da publicação, constava também toda a Divisão Territorial do Brasil.

A publicação apresentou as informações para os 5 507 municípios criados e instalados em 1º de agosto de 2000 e, em anexo, para os 54 novos municípios, que foram instalados em 1º de janeiro de 2001.

12.2 Sinopse Preliminar

Dando continuidade à divulgação dos resultados preliminares do Censo 2000, foi apresentada, em maio de 2001, a *Sinopse preliminar do censo demográfico 2000*, dando seguimento à série histórica desta publicação, iniciada com o Censo de 1940.

Assim como na publicação dos resultados preliminares, divulgada em dezembro de 2000, os dados da sinopse preliminar tiveram como fonte o SIGC. Foram considerados para divulgação os dados atualizados no sistema até 31 de janeiro de 2001, ou seja, já levando em consideração algumas mudanças decorrentes da atividade de reabertura da coleta de setores em algumas unidades da federação (ver 9.6 Evolução da coleta). Assim, as informações da sinopse também diferiram dos dados definitivos do censo, divulgados mais tarde.

A publicação, além das notas metodológicas, apresentou uma ampla retrospectiva dos dados dos censos desde 1872, acompanhada de textos analíticos sobre a dinâmica da população brasileira e sobre os domicílios.

No que toca ao plano tabular, foram apresentados: os dados sobre população residente, segundo o sexo e situação do domicílio; e domicílios, segundo a espécie, para as grandes regiões, unidades da federação, regiões metropolitanas e todos os municípios do país. No anexo da publicação, encontram-se informações para os 54 novos municípios, instalados em 1º de janeiro de 2001.

No CD-ROM, que acompanhou a publicação impressa, foram reunidas todas as tabelas desse volume; entretanto, a divulgação só atingiu o nível de distrito.

12.3 Resultados do Conjunto Universo

Antes do início do trabalho de crítica e imputação dos dados, foi necessário fazer a formação do Conjunto Universo, a partir das informações comuns do Questionário Básico e do Questionário da Amostra.

A atividade de crítica teve início com a formação dos lotes de trabalho, tendo seguimento com a definição das funções de crítica, que analisaram, em separado, as inconsistências das informações relativas aos domicílios e às pessoas. Para a análise dos dados de pessoa, foi necessário implementar as etapas da crítica intra e entre registros, onde, respectivamente, a investigação e correção dos erros levava em consideração variáveis para uma ou entre pessoa(s) moradora(s) de um mesmo domicílio.

12.3.1 Formação do Conjunto Universo

A formação do Conjunto Universo do Censo Demográfico 2000 consistiu na reunião dos domicílios e das pessoas investigados no Questionário Básico – CD 1.01 – e no Questionário da Amostra – CD 1.02 – associando a cada um(a) deles(as) o conjunto de informações comuns aos dois questionários, ou seja, aquelas coletadas para 100% da população.

As variáveis comuns aos dois questionários são:

a) variáveis de domicílio

- Espécie ou Espécie do domicílio – V0201;
- Tipo ou Tipo de domicílio – V0202;
- Condição de ocupação do domicílio – V0203 no CD 1.01 e V0205 no CD 1.02;
- Condição de ocupação do terreno do domicílio – V0204 no CD 1.01 e V0206 no CD 1.02;
- Forma de abastecimento de água – V0205 no CD 1.01 e V0207 no CD 1.02;
- Canalização da água – V0206 no CD 1.01 e V0208 no CD 1.02;
- Número de banheiros – V0207 no CD 1.01 e V0209 no CD 1.02;
- Existência de sanitário – V0208 no CD 1.01 e V0210 no CD 1.02;
- Tipo de escoadouro – V0209 no CD 1.01 e V0211 no CD 1.02;
- Destino do lixo – V0210 no CD 1.01 e V0212 no CD 1.02;

b) variáveis de pessoa

- Sexo - V0401;
- Relação com a pessoa responsável pelo domicílio - V0402;
- Mês e ano de nascimento – V0403 no CD 1.01 e V0405 no CD 1.02;
- Idade em 31 de Julho de 2000 – V0404 no CD 1.01 e V0406 no CD 1.02;

- Idade presumida – V0405 no CD1.01 e V0407 no CD1.02;

- Sabe ler e escrever – V0406 no CD1.01 e V0428 no CD1.02;

- Curso mais elevado que freqüentou no qual concluiu pelo menos uma série[1]

V0407 no CD 1.01 ou
V0432 no CD 1.02, para a pessoa que não estava freqüentando escola, mas já havia freqüentado; e
V0430 no CD 1.02, para a pessoa que estava freqüentando escola a partir da segunda série; ou o grau do curso imediatamente anterior ao declarado no quesito 4.30, para a pessoa que estava freqüentando escola na primeira série.

- Última série concluída com aprovação[1] – V0408 no CD 1.01 ou

V0433 no CD 1.02, para a pessoa que não estava freqüentando escola, mas já havia freqüentado; e
V0431 no CD 1.02, para a pessoa que estava freqüentando escola a partir da segunda série; ou a última série correspondente ao grau do curso imediatamente anterior ao declarado no quesito 4.30, para a pessoa que estava freqüentando escola na primeira série.

- Rendimento bruto do mês de Julho de 2000[1] – valor declarado no quesito 4.09 no CD 1.01, e o somatório dos registrados nos seguintes quesitos do CD 1.02:

Rendimento no trabalho principal – 4.51;

Rendimento nos demais trabalhos – 4.52;

Proventos de aposentadoria ou pensão – 4.57;

Proventos de aluguel – 4.58;

Proventos de pensão alimentícia, mesada ou doação recebida de não-morador – 4.59;

Proventos de renda mínima, bolsa escola, etc. – 4.60; e

Proventos de outros rendimentos – 4.61.

12.3.2 Crítica e Imputação dos Dados

O desenvolvimento da crítica e imputação das informações do Conjunto Universo tiveram início com a definição dos lotes de apuração que, em última análise, constituíram as bases de dados a serem submetidas aos aplicativos de depuração das inconsistências.

Para o trabalho de detecção e correção das incompatibilidades dos dados do Conjunto Universo foi utilizado o sistema DIA - Detección e Imputación Automática de errores para datos cualitativos, que se acha descrito, de maneira breve, no anexo de CD-ROM desse capítulo[2]. Como o DIA não admite a

[1] No Conjunto Universo, essa variável apresenta valor somente para a Pessoa Responsável pelo Domicílio ou Individual em Domicílio Coletivo

[2] Caso o leitor não conheça o sistema DIA, recomenda-se consultar o anexo, antes de dar seqüência à leitura do capítulo.

utilização de funções de crítica que envolvam variáveis de registros distintos, foi necessário implementar uma estratégia que permitisse superar essa limitação do sistema.

Foram constituídos lotes de trabalho e, em cada um, a seqüência de execução dos aplicativos através do sistema DIA foi a seguinte: Características da Pessoa, que tratava as críticas entre registros; e, para as críticas intra registros, Características do Domicílio, Características da Pessoa Responsável pelo Domicílio ou Individual em Domicílio Coletivo, e Características das Demais Pessoas.

12.3.2.1 Formação dos lotes

Assim que os dados do SIGC estavam disponíbilizados, estabeleceram-se critérios para a formação dos lotes de apuração, com vistas à execução da detecção e correção automática dos erros através do sistema DIA. Esses critérios foram os mesmos utilizados no Censo Demográfico 1991.

Um fator importante para a definição dos critérios de formação dos lotes é a proximidade geográfica, o que significa ter os questionários de uma mesma região geográfica em um mesmo lote, o que constitui-se num fator de homogeneidade de características. Essa homogeneidade é importante em função da metodologia de imputação utilizada, que se baseia na distribuição dos dados dos registros "bons" observados no lote, entendendo-se como tal aqueles que não apresentam qualquer erro, segundo as regras de crítica definidas.

Os critérios para a formação dos lotes basearam-se nos seguintes pontos:

a) obtenção do menor número possível de lotes, para minimizar o número de relatórios a serem analisados, após cada aplicação do sistema DIA;

b) obtenção de um tamanho mínimo a fim de viabilizar o processo de correção, tendo em vista a utilização das distribuições de registros "bons", como base da imputação; e

c) geração dos lotes levando em conta a situação do domicílio (urbana e rural), bem como a divisão geográfica do país, contemplando as partições de cada unidade da federação, através da ordenação dos respectivos setores segundo a mesorregião, microrregião, município, distrito e subdistrito.

A quantidade de domicílios particulares ocupados – DPO – fornecida pelo SIGC - foi o ponto de partida para o processo de formação dos lotes. A escolha de um tamanho máximo de 90 000 domicílios para cada lote, foi feita levando em conta um acréscimo ao número estabelecido para o Censo de 1991, que foi de 70 000.

A quantidade de lotes a serem formados, segundo a situação do domicílio, em cada unidade da federação, foi o resultado da divisão do número de domicílios particulares ocupados por 90 000, arredondado para o inteiro seguinte.

O tamanho, aproximado, de cada lote em cada unidade da federação, por situação do domicílio, foi obtido pela divisão do correspondente DPO pelo número de lotes encontrado. De posse desse tamanho aproximado, cada lote foi formado, fazendo-se os cortes na relação ordenada de setores citada no item "c".

Baseando-se nesses critérios, foram gerados 526 lotes de apuração, sendo 429 urbanos e 97 rurais, que vão apresentados na tabela seguinte, acompanhados do número de domicílios particulares ocupados, informados no SIGC, por unidade da federação e situação do domicílio.

Tabela 12.1 - Número de lotes e quantidade de domicílios particulares ocupados no SIGC, do Conjunto Universo, por situação do domicílio, segundo as Unidades da Federação

Unidades da Federação	Situação do domicílio			
	Urbana		Rural	
	Domicílios particulares ocupados	Número de lotes	Domicílios particulares ocupados	Número de lotes
Brasil	**37 455 153**	**429**	**7 567 163**	**97**
Rondônia	229 944	3	121 533	2
Acre	91 093	2	39 619	1
Amazonas	455 707	6	123 199	2
Roraima	59 368	1	16 135	1
Pará	913 540	11	411 420	5
Amapá	89 387	1	10 100	1
Tocantins	212 501	3	70 667	1
Maranhão	757 926	9	484 191	6
Piauí	430 566	5	233 103	3
Ceará	1 294 941	15	468 621	6
Rio Grande do Norte	505 278	6	168 565	2
Paraíba	624 314	7	227 840	3
Pernambuco	1 558 239	18	420 682	5
Alagoas	463 455	6	192 227	3
Sergipe	320 653	4	118 685	2
Bahia	2 218 482	25	976 917	11
Minas Gerais	3 977 365	45	805 740	9
Espírito Santo	685 193	8	160 242	2
Rio de Janeiro	4 107 268	46	157 028	2
São Paulo	9 756 179	109	639 303	8
Paraná	2 216 678	25	464 969	6
Santa Catarina	1 205 879	14	298 814	4
Rio Grande do Sul	2 518 408	28	534 051	6
Mato Grosso do Sul	480 028	6	89 406	1
Mato Grosso	525 143	6	134 112	2
Goiás	1 231 918	14	177 167	2
Distrito Federal	525 700	6	22 827	1

Fonte: IBGE, Censo Demográfico 2000, Sistema de Indicadores Gerenciais da Coleta.

12.3.2.2 Tratamento das omissões da variável "espécie do domicílio"

Uma forma de resolver essa questão, é fazer a imputação dos valores em branco dessa variável – V0201 – levando em conta o preenchimento ou não da seqüência dos demais campos do bloco Características do Domicílio. No entanto, preferiu-se não adotar essa estratégia, pela existência de erros de preenchimento que trariam prejuízo à imputação.

Assim, durante a crítica intra-registros, quando da execução do Aplicativo Características do Domicílio, essa variável foi tratada como fixa no sistema DIA, o que recomendava um tratamento prévio que eliminasse as omissões de informação. Resolveu-se, então, considerar o preenchimento do quesito 1.09 – Número na Folha de Domicílio Coletivo, pertencente ao bloco Identificação. A solu-

ção implicou na criação da variável auxiliar V1090 “existência de domicílio coletivo”que classificava o domicílio em particular ou coletivo, respectivamente, quando a V0109 assumisse o valor zero ou outro qualquer.

Assim, a detecção de omissão e a conseqüente imputação determinística para a variável V0201 foi feita através de procedimento específico, implementado durante o processo de formação dos lotes de trabalho a serem submetidos aos aplicativos do DIA, da seguinte forma:

- V0201 era igual a 1 – Particular permanente –, quando a V1090 indicasse a não-existência de domicílio coletivo; e .

- V0201 era igual a 3 – Coletivo –, quando a V1090 indicasse o contrário.

12.3.2.3 Tratamento das omissões da variável "sexo"

A princípio, não estava previsto fazer o tratamento prévio das situações de omissão da variável “sexo” V0401. Assim, tentou-se tratar essa situação realizando sua depuração de acordo com as estratégias definidas para o aplicativo Características da Pessoa – na crítica entre registros – ou seja, em conjunto com as demais inconsistências. Entretanto, a análise de alguns dos resultados da imputação mostrou a inconveniência da utilização desse procedimento.

A solução encontrada foi executar um aplicativo DIA, apenas para a correção das omissões na variável “sexo”. A estratégia elaborada para esse aplicativo foi fazer a imputação através de distribuição conjunta, baseada na variável auxiliar criada V4702 “grupo qüinqüenal de idade” e na variável V0402 “relação com a pessoa responsável pelo domicílio” através do método proporcional.

É importante ressaltar que a variável “sexo” poderia, durante a execução do aplicativo seguinte, Características da Pessoa, sofrer nova alteração, caso a categoria que lhe fora atribuída ficasse inconsistente perante o conjunto estabelecido para as funções de crítica entre registros.

12.3.2.4 Crítica entre registros

A crítica dos dados do Conjunto Universo que levava em conta as regras de crítica entre registros, foi realizada pelo Aplicativo Características da Pessoa, sendo apenas objeto da imputação as variáveis V0401 e V0402.

Inicialmente, para que os lotes pudessem ser submetidos ao sistema de crítica foi necessária a execução de um programa de ordenação lógica das pessoas em cada domicílio. Os critérios para essa ordenação foram definidos levando-se em conta os procedimentos estabelecidos no Manual do Recenseador para a elaboração da lista de moradores e a idade das pessoas; no documento *Esquema de ordenação lógica das pessoas no questionário básico - censo 2000* (2001), podem ser consultados mais detalhes desse trabalho.

Para que a crítica entre registros pudesse ser executada através do DIA, foi necessária uma estratégia especial de criação de um novo arquivo, onde as informações das pessoas moradoras de um mesmo domicílio foram rearrumadas, de forma a comporem um único registro, considerando-se todos os domicílios com até quarenta moradores.

A variável V0402 foi imputada através da distribuição condicional, que levava em conta a variável auxiliar V0702 “grupo de idade da pessoa” através do método proporcional.

Embora durante a execução da crítica entre registros o total de pessoas do domicílio estivesse correto, pois a informação já havia sido tratada pela crítica quantitativa nos Centros de Captura de Dados – CCDs, os totais por sexo poderiam sofrer alterações em razão das imputações realizadas. Em razão disso, após a execução do DIA, os totais por sexo tiveram que ser recalculados.

A existência de domicílios com mais de quarenta moradores pôde ser constatada nas seguintes Unidades da Federação: Rio Grande do Sul; Mato Grosso; Pará; Pernambuco; Minas Gerais; São Paulo; Sergipe e Ceará. Para cada uma dessas unidades, encontrou-se apenas um único domicílio cujo número de moradores era, respectivamente: 49; 62; 64; 42; 54; 54; 43 e 59. Dessas situações, em quatro unidades – RS, MT, PA e MG – detectou-se inconsistências na crítica entre registros; esses casos foram corrigidos manualmente.

12.3.2.5 Crítica intra-registros

As condições de imputação dos aplicativos do sistema DIA, para as funções de crítica intra-registros dos dados do Conjunto Universo, são apresentadas a seguir.

a) Aplicativo Características do Domicílio

Neste aplicativo poderiam ser imputadas todas as variáveis do bloco 2 – Características do Domicílio - com exceção da V0201 "espécie do domicílio", já consistente, conforme explicado anteriormente e, portanto, mantida fixa durante a execução do aplicativo.

As variáveis foram imputadas de acordo com as respectivas distribuições marginais formadas pelas freqüências dos registros não suspeitos, através do método proporcional.

Para as variáveis V0203 "condição de ocupação do domicílio" e V0207 "número de banheiros" foram atribuídos pesos 2 e 1, respectivamente, diferentemente das demais, cujos pesos foram mantidos em 5, peso médio da escala de confiança na variável. O motivo para essa alteração decorreu dos resultados das análises efetuadas, onde se constatou leve mudança em algumas distribuições dessas variáveis antes e depois da imputação.

b) Aplicativo Características da Pessoa Responsável pelo Domicílio ou Individual em Domicílio Coletivo

Neste aplicativo podiam ser imputadas as seguintes variáveis: idade, sabe ler e escrever, curso mais elevado que freqüentou, no qual concluiu, pelo menos uma série e a última série concluída com aprovação.

Considerando-se que a "idade" pode ser obtida através do mês e ano de nascimento, ou da idade em 31 de julho de 2000, ou ainda, da idade presumida, foi necessário criar um algoritmo que, levando em consideração critérios para esses três quesitos, chegava a informação da idade a ser tratada nesse aplicativo. Embora o algoritmo fizesse a escolha, era possível que a informação da idade passasse por correção, durante a execução do DIA, visto que o sistema poderia identificar inconsistências de acordo com as regras de crítica em que esta variável estivesse envolvida.

Historicamente, admite-se que a informação sobre a idade constitua-se num dado com elevado grau de confiança. Por esse motivo, como critério de imputação, atribuiu-se-lhe peso 1, enquanto que as demais variáveis tiveram o peso médio 5, garantindo-se, desse modo, que a idade fosse, em relação às demais, proporcionalmente, bem menos imputada.

Para a imputação da "idade", utilizou-se a distribuição conjunta, a partir da criação da variável auxiliar V4040 "faixa de idade do cônjuge" e da variável V0402.

A variável "sabe ler e escrever" foi imputada de acordo com a distribuição conjunta com a variável "idade", utilizando-se a distribuição dos registros não-suspeitos e o método proporcional.

As variáveis "curso mais elevado que freqüentou, no qual concluiu, pelo menos uma série" e "última série concluída com aprovação", foram imputadas pelo DIA, de acordo com a situação, através de método determinístico ou probabilístico. A imputação determinística passou a ser uma estratégia em virtude das eventuais inconsistências oriundas de informações errôneas entre a série e o grau, envolvendo mudanças no sistema de ensino brasileiro, ao longo do tempo. Nos casos em que a estratégia foi a imputação probabilística, utilizou-se a distribuição marginal dos registros não-suspeitos e o método proporcional.

c) Aplicativo Características das Demais Pessoas

Os procedimentos utilizados para a execução deste aplicativo foram os mesmos já descritos para o aplicativo anterior. Houve, apenas, a necessidade de se criar a variável auxiliar V4041 "faixa de idade da pessoa responsável ou individual em domicílio coletivo" para imputação da idade das demais pessoas.

12.3.2.6 Análise do processo de crítica e imputação

No Censo Demográfico 2000, os procedimentos de crítica e imputação dos dados foram constantemente monitorados a fim de evitar a alteração na estrutura da informação. Vários foram os instrumentos utilizados com esse objetivo, como as tabelas (conjunto de tabelas que envolvem o aplicativo), a análise demográfica, estudos de população e o controle das alterações nas respostas originais constantes do questionário.

a) Análise dos relatórios do Sistema DIA

O trabalho de análise do processo de crítica e imputação dos Resultados do Universo foi desenvolvido, para a crítica entre registros e para cada um dos aplicativos da crítica intra registros, em duas partes: a que permitia avaliar a correção automática dos erros detectados em nível de cada lote de trabalho e a outra, com o mesmo objetivo, abrangendo os municípios e alguns subdistritos selecionados. Os elementos para a realização dessa tarefa constam do Plano de Análise da Correção Automática e Elementos de Apoio para a Análise da Composição do Lote – CD 1.01 – Questionário Básico.

Obedecendo às determinações do plano, para a investigação em nível de lote eram emitidos relatórios que apresentavam dados gerais sobre o resultado da imputação, informando, por exemplo:

- os totais de registros, de registros bons e de registros com erros, em valores absolutos e relativos;

- a participação de cada tipo de erro em relação ao total de registros;

- os registros segundo o tipo de imputação, apresentando variáveis com valores inválidos, com inconsistências entre variáveis; e

- número de variáveis imputadas por número de registros.

Além dessas informações, toda vez que um lote era considerado suspeito em algum aplicativo, era também emitido o relatório Tablas, parte integrante do sistema DIA. Através desse relatório, era possível analisar, para cada variável, as distribuições de freqüências de entrada e de saída dos dados, assim como, as distribuições dos registros bons e os não-suspeitos, procurando identificar distorções significativas resultantes do processo de imputação.

Um lote era considerado suspeito quando apresentasse alguma variável fixa com valor inválido, ou atingisse o limite de tolerância estabelecido em, pelo menos, um dos seguintes indicadores;

- E - percentual de registros com erro em relação ao total de registros; e

- F - percentual de registros que falharam em cada regra de crítica em relação ao total de registros com erro.

Esses indicadores, calculados após a imputação para cada um dos aplicativos da crítica intra-registros, tinham como limites máximos 10% e 50%, respectivamente, para E e F.

No entanto, era necessário estabelecer um outro nível de investigação que permitisse uma análise mais desagregada dos registros, de modo a possibilitar a identificação de eventuais distorções proporcionadas pela imputação, não-sensíveis no nível agregado de lote.

Desenvolveu-se, então, no plano de análise, os critérios para a emissão de relatórios que permitissem efetuar a investigação para municípios ou subdistritos suspeitos. Isso permitiu realizar o trabalho em dimensão bem próxima ao das análises estruturais desenvolvidas pelos especialistas das diversas áreas temáticas da DPE.

Um município ou subdistrito foi considerado suspeito, caso alcançasse os limites de tolerância para, pelo menos, um dos indicadores seguintes:

E - já definido anteriormente;

Ij - percentual de registros em que a variável j apresentou valor inválido em relação ao total de registros;

$\max_i D_j(i)$ - maior distância em termos relativos, entre as freqüências marginais dos dados bons (FB) e dos dados depurados (FD), para o código i (valores possíveis) da variável j, onde:

$$D_j(i) = \left| \frac{FD_j(i)}{FD_j} - \frac{FB_j(i)}{FB_j} \right| x100$$

FD = nº total de registros depurados da variável j

FB = nº total de registros bons da variável j

$\max_i A_j(i)$ - maior distância em termos relativos, entre as freqüências marginais dos dados de entrada (FE) e dos dados depurados (FD), para o código i da variável j, onde:

$$A_j(i) = \left| \frac{FD_j(i) - FE_j(i)}{FD_j} \right| x100$$

FD = nº total de registros depurados da variável j

FE = nº total de registros de entrada da variável j

Tj - distância entre as freqüências dos dados bons (FB) e os dados depurados (FD), para a variável j, em termos relativos, sendo n o número de códigos possíveis para a variável j, onde:

$$T_j = \sum_{i=1}^{n} \frac{FD_i - FB_i}{\frac{FD}{2}}$$

Bj - distância entre as freqüências dos dados de entrada (FE) e dos dados depurados (FD) para a variável j, em termos relativos, sendo n o número de códigos possíveis para a variável j, onde:

$$B_j = \sum_{i=1}^{n} \frac{FD_i - FE_i}{\frac{FD}{2}}$$

O indicador Ij foi somente calculado para os aplicativos "Características da Pessoa Responsável pelo Domicílio ou Individual em Domicílio Coletivo" e "Características das Demais Pessoas". Quanto aos limites de tolerância, foram estabelecidos 10% e 5% para, respectivamente, E e Bj e 3% para os demais indicadores.

Tanto para o lote quanto para o município ou subdistrito, fundamentalmente o processo de avaliação da imputação através das tablas, concentrava-se na análise dos registros de entrada e depurados, procurando-se identificar alterações nas distribuições das variáveis. Como suporte a esse trabalho, caso necessário, eram consultadas também as informações sobre a composição do lote da crítica.

Qualquer problema – nos dados gerais ou nas tablas – encontrado no resultado da imputação, era encaminhado aos analistas temáticos para o exame da consistência das variáveis correspondentes ao aplicativo implementado.

b) Análise da consistência da imputação e validação dos resultados

É importante fazer uma primeira observação com respeito à Análise do processo de Crítica e Imputação e à dificuldade em separá-la do item validação dos resultados.

Procederam-se análises para algumas unidades da federação e, em alguns casos, alguns municípios onde era estudado o comportamento de cada quesito do questionário antes e depois da imputação automática. Para isso, foram utilizados a visualização de imagens, a listagem do registro completo das pessoas contendo as variáveis a serem analisadas e matrizes de contingência conforme modelo a seguir.

Figura 12.1 - Matriz de contingência com vetores antes e após o processo de imputação

VPOSTERIOR / VANTERIOR	Total 1	%	1	2	3	...	j	...	n
Total2									
%									
1									
2									
3									
...									
j									
...									
m									

Onde:

VANTERIOR = código do quesito antes da correção automática

VPOSTERIOR = código do quesito após a correção automática

A acumulação de valores na diagonal indica ausência de modificações no processo, portanto constituíram o alvo do estudo os casos em que houve significativo aumento, ou redução, de frequência de casos observados em alguma categoria fora da diagonal. As distorções, em geral, desapareceram com a revisão e alteração de algumas regras contidas no plano de crítica. Em alguns casos elas estavam justificadas por se tratar de correção de erros sistemáticos.

As listagens dos registros e as matrizes de contingência foram obtidas através da utilização do REDATAM+G4 (REcuperação de DAdos para Áreas pequenas por Microcomputador, 4ª Geração), um programa computacional desenvolvido pelo Centro Latino-americano e Caribenho de Demografia – CELADE. Com este objetivo foi feita a junção dos arquivos antes e depois da imputação.

Variáveis de domicílio

Na fase de crítica e imputação dos dados de domicílios do Conjunto Universo do Censo Demográfico 2000 foram realizadas análises de consistência dos resultados de cada característica investigada em relação aos obtidos no Censo Demográfico 1991, utilizando indicadores da Pesquisa Nacional por Amostra de Domicílios na década de 1990 como balizamento de tendência. Buscou-se verificar, também, a ocorrência de efeitos de registros inadequados que pudessem ter acontecido na etapa de coleta ou de crítica e imputação dos dados, considerando,

separadamente, os resultados obtidos por meio dos Questionários Básico e da Amostra para as parcelas urbana e rural e, ainda, o cruzamento de determinadas características investigadas. Com base nessas análises constatou-se que, a maior incidência de falhas oriundas da fase de coleta foram decorrentes da desobediência da seqüência dos quesitos nos casos em que a rota era definida em função das respostas registradas. Constatou-se, também, que apenas nos quesitos que geravam seqüências distintas em função das respostas registradas e para aqueles que deveriam ser seguidos somente em função de determinada respostas registradas no anterior, o procedimento geral de crítica e imputação aplicado para correção dos erros de seqüência e dos registros omitidos apresentava efeito perceptível na distribuição dos resultados dos itens. Para evitar esta ocorrência, foram adotados procedimentos especiais para correção e imputação, existentes no próprio sistema DIA, conforme descritos no item 12.3.2.5 a) Aplicativo Características do Domicílio.

Sexo e idade

Para validar o grau de precisão dos resultados das declarações de sexo e de idade dos entrevistados, calcularam-se alguns indicadores demográficos, objetivando verificar as imperfeições nas declarações e se as informações eram coerentes com a tendência observada ao longo dos censos. Os métodos de Myers, Bachi e Whipple (SHRYOCK et al., 1971) foram utilizados para avaliar o grau de atração e repulsão exercido pelos dígitos terminais.

Os índices foram calculados com as informações originais do campo, sem aplicação da correção da crítica qualitativa e com as informações corrigidas, quer dizer, após a passagem da correção automática para o ano 2000 e comparada com os censos a partir de 1940.

Índices preferenciais calculados

Uma forma de comprovar a coerência interna das respostas é verificar a tendência dos informantes em declarar determinados dígitos terminais para a idade. Geralmente é comum o homem declarar-se com 21 anos, porque corresponde à maioridade, enquanto que as mulheres procuram reduzir a idade. As informações errôneas também podem ser causadas por razões econômicas, sociais, políticas ou puramente individuais. Normalmente, existe uma tendência a arredondar a idade, acumulando-se, portanto declarações em idades terminadas em 0 e 5 anos.

O Índice de Myers pode assumir valores entre zero e cento e oitenta correspondendo, respectivamente, a informações de idade prestadas com exatidão e a todas as declarações de idades terminadas pelo mesmo dígito. Os índices calculados com os resultados do Censo Demográfico 2000 não revelaram qualquer variação entre as informações originais, sem tratamento de crítica e as corrigidas, após a passagem da correção automática. E, quando cotejados com o Censo Demográfico 1991, revelaram que passou a existir um ligeiro crescimento na atração por determinados dígitos terminais. Isto é natural, considerando que o Censo Demográfico 1991 foi o único que não foi realizado em ano finalizado em zero. Na análise por sexo, contrariando uma tendência dos censos, o menor grau de precisão da declaração de idade foi proveniente das informações dos homens.

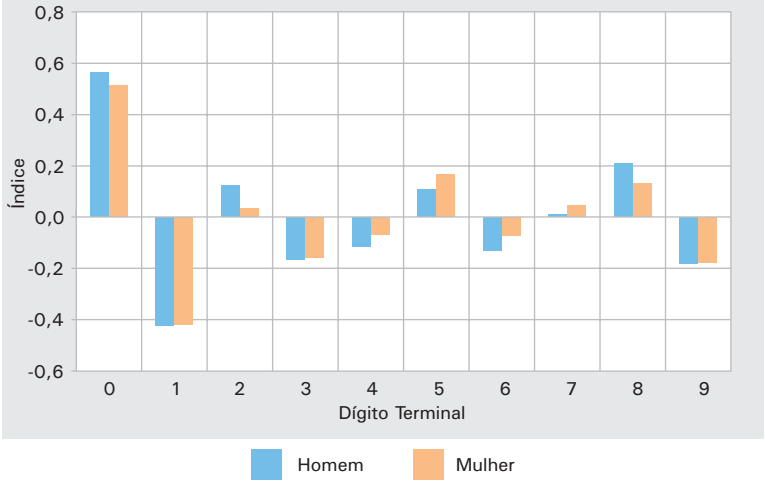
Tabela 12.2 - Índices preferenciais calculados para a população residente
Brasil

Ano	Índices preferenciais calculados		
	Myers	Bachi	Whipple
1940	17,9	12	148
1950 (1)	14,4	9,6	140
1960 (1)	17,5	11,1	143
1970	10,2	6,6	128
1980	4,1	2,6	111
1991	1,3	0,8	103
2000			
Informações originais	1,9	1	104
Informações corrigidas	1,9	1	104

Fonte: IBGE, Censo Demográfico 1940/2000.
(1) População presente.

A preferência por alguns dígitos terminais, no Censo Demográfico 2000 foi especialmente pelo zero, seguido pelos dígitos oito e cinco, e a repulsão foi pelo um, seguido pelo nove. O ano do levantamento apresenta uma certa influência na idade, porque nos censos terminados em zero a tendência é de dígito terminal atrativo zero e repulsivo um (1). No Censo Demográfico 1991, o dígito terminal preferido foi especialmente o cinco e o repulsivo foi o nove.

Gráfico 12.1 - Índice de Myers com as informações corrigidas



O Índice de Whipple tem como objetivo medir a concentração das declarações nas idades terminadas pelos dígitos zero e cinco. O índice de Whipple é o quociente entre duas distribuições de população. O numerador corresponde ao somatório do número de pessoas nas idades pontuais terminadas em zero e cinco a partir de 25 anos até 60 anos de idade multiplicada por cinco, assumindo a hipótese de linearidade no intervalo, e o denominador corresponde ao somatório das pessoas no intervalo de idade entre 23 e 62 anos.

$$IW = [\, 5\, [P(25)+P(30)+P(35)+.....+P(55)+P(60)]\, x\, 100] \,/\, {}^{62}\Sigma_{x=23}\, P(x)$$

onde,

P(x) = população na idade x.

A classificação utilizada pelo *Demographic yearbook* é a seguinte: IW < 105 correspondem a dados muitos precisos e IW > 175 correspondem a dados pouco refinados. Como o resultado foi IW = 104, considerou-se que os dados são precisos e não apresentaram diferença entre os sexos. Em geral existe a tendência à declaração da idade em certos números, especialmente nos que terminam em zero ou cinco, seja porque os entrevistados não conhecem exatamente sua idade ou porque não compreendem a importância de declarar a idade exata.

O método de Bachi aplica o método de Whipple repetidamente para determinar a extensão de preferência para cada dígito final e a partir daí baseia-se, tal como o Índice de Myers, na soma dos desvios, tomados positivamente, entre a freqüência relativa com que cada dígito de zero a nove ocorreu em um determinado levantamento e a freqüência relativa esperada, caso não houvesse preferência por nenhum dígito (distribuição uniforme). Os resultados obtidos pelo método de Bachi se assemelham aos obtidos pelo método de Myers.

As distorções e as falhas nas declarações de sexo e idade são menores de um censo para outro, em função da diminuição da dificuldade da população em informar sua idade com precisão, o que reflete positivamente na qualidade dos diversos indicadores. Os índices atingiram magnitude tão baixa que as oscilações podem ser consideradas desprezíveis.

Outros indicadores também foram calculados para avaliar as possíveis distorções da estrutura por sexo e idade da população, tais como as pirâmides etárias, razões de masculinidade, as razões de idade dentre outros.

Sexo ignorado

Em um primeiro momento, o critério para imputação do sexo, nos casos em que esta informação não foi coletada, levou em consideração somente a observação da variável "relação com o responsável pelo domicílio". Este procedimento não apresentou resultados satisfatórios, uma vez que levou à imputação, em maior quantidade, de sexo masculino nas idades mais avançadas, pois as pessoas envolvidas eram, na sua maioria, responsáveis por domicílio. O resultado não condizia com o esperado, que seria uma proporção maior de mulheres nessas referidas idades.

Foi observado que a ausência de declarações de sexo se concentrava nas primeiras idades e nas mais avançadas. Então, foi solicitada a inclusão da variável "grupos de idade qüinqüenais" como condicionante do sexo a ser imputado – ver item 12.3.2.3. Dessa forma, se o sexo ignorado fosse observado numa pessoa com idade mais avançada, haveria uma maior probabilidade de ser imputado sexo feminino, caso contrário, masculino. Por outro lado, nas primeiras idades a probabilidade de ser imputado o sexo masculino seria maior que a de ser imputado feminino.

Relação com o responsável pelo domicílio

Em alguns questionários de domicílios particulares, o quesito "Qual é a relação com o responsável pelo domicílio?" havia sido preenchido incorretamente. Ao invés de ter sido assinalada, para a primeira pessoa, a quadrícula "pessoa responsável", havia sido assinalada a quadrícula "individual em domicílio coletivo".

De acordo com a ordem lógica estabelecida no manual de crítica, esta pessoa passaria à posição de outro membro do domicílio, enquanto a pessoa que ocupasse a segunda posição no questionário passaria à primeira pessoa e, consequentemente, a ser a responsável pelo domicílio.

Fazendo-se um estudo das características das pessoa que ocupavam a primeira posição do questionário, comparando-a com as demais pessoas e considerando a estrutura domiciliar, observou-se que tratava-se de erro de preenchimento da informação referente à responsabilidade pelo domicílio.

Com o objetivo de que não fosse alterada a estrutura domiciliar, foi acrescido aos critérios de ordem lógica, a manutenção dessa pessoa como primeira pessoa moradora no domicílio. Assim os responsáveis com a quadrícula "individual em domicílio coletivo" indevidamente assinalada, sofreriam acerto nesta informação quando da consistência feita pelo sistema DIA.

Freqüências de imputação

As tabelas 12.3 a 12.6 mostram a freqüência de imputações das variáveis do Conjunto Universo, relativas a domicílios e pessoas. Apresentam também o número de domicílios e pessoas, que foram objeto de imputação.

Tabela 12.3 - Registros imputados, segundo as variáveis de domicílio - Brasil

Variável	Total	Sem imputação		Com imputação	
		Absoluto	Relativo (%)	Absoluto	Relativo (%)
V0201	45 507 516	45 337 228	99,63	170 288	0,37
V0202	45 507 516	44 920 635	98,71	586 881	1,29
V0203	45 507 516	44 906 225	98,68	601 291	1,32
V0204	45 507 516	44 651 605	98,12	855 911	1,88
V0205	45 507 516	45 121 014	99,15	386 502	0,85
V0206	45 507 516	45 044 603	98,98	462 913	1,02
V0207	45 507 516	45 204 565	99,33	302 951	0,67
V0208	45 507 516	44 942 066	98,76	565 450	1,24
V0209	45 507 516	45 260 342	99,46	247 174	0,54
V0210	45 507 516	45 263 848	99,46	243 668	0,54

Fonte: IBGE, Censo Demográfico 2000.

Tabela 12.4 - Imputação nos registros de domicílio, segundo o aplicativo - Brasil

Aplicativo	Absoluto	Relativo (%)
Total	**45 507 516**	**100,00%**
Sem imputação	42 348 158	93,06%
Imputação somente na V0201	52 568	0,12%
Imputação pelo Aplicativo "Características do Domicílio" (V0202 a V0210)	2 989 070	6,57%
Imputações V0201 e pelo Aplicativo "Características do Domicílio" (V0202 a V0210)	117 720	0,26%

Fonte: IBGE, Censo Demográfico 2000.

Tabela 12.5 - Registros imputados, segundo as variáveis de pessoa - Brasil

Variável	Total	Sem imputação		Com imputação	
		Absoluto	Relativo (%)	Absoluto	Relativo (%)
V0401	169 799 170	167 980 331	98,93%	1 818 839	1,07%
V0402	169 799 170	168 434 764	99,20%	1 364 406	0,80%
V4322	169 799 170	169 620 079	99,89%	179 091	0,11%
V4344	169 799 170	169 794 559	100,00%	4 611	0,00%
V0406	169 799 170	167 643 042	98,73%	2 156 128	1,27%
V0407	45 507 516	45 170 336	99,26%	337 180	0,74%
V0408	45 507 516	45 328 064	99,61%	179 452	0,39%
V4093	45 507 516	44 701 199	98,23%	806 317	1,77%

Fonte: IBGE, Censo Demográfico 2000.

Nota: As variáveis V0407, V0408 e V4093 só foram investigadas para a pessoa responsável pelo domicílio ou individual em domicílio coletivo.

Tabela 12.6 - Imputação nos registros de pessoa, segundo o aplicativo - Brasil

Aplicativo	Absoluto	Relativo (%)
Total	**169 799 170**	**100,00**
Sem imputação	159 649 188	94,02
Imputação pelo Aplicativo "Características da Pessoa"	1 405 900	0,83
Imputação pelo Aplicativo "Características da Pessoa Responsável pelo Domicílio ou Individual em Domicílio Coletivo"	6 842 666	4,03
Imputação pelos Aplicativos "Características da Pessoa" e "Características da Pessoa Responsável pelo Domicílio ou Individual em Domicílio Coletivo"	122 251	0,07
Imputação pelo Aplicativo "Características das Demais Pes	1 709 415	1,01
Imputação pelos Aplicativos "Características da Pessoa" e "Características das Demais Pessoas"	69 750	0,04

Fonte: IBGE, Censo Demográfico 2000.

12.3.2.7 Imputação da Variável de Rendimento

Este item descreve o processo de imputação da variável de rendimento dos responsáveis por domicílios do conjunto universo do Censo Demográfico 2000. São apresentadas as motivações que levaram ao desenvolvimento de tal processo, bem como a metodologia desenvolvida para sua aplicação e os resultados obtidos.

Vale lembrar que estamos tratando do Conjunto Universo, cuja formação está descrita no início deste capítulo.

Dentre as perguntas aplicadas, havia a que indagava o valor do "rendimento bruto do mês de julho de 2000" (em R$), proveniente de trabalho e de outras fontes, obtido pela pessoa responsável pelo domicílio ou pelo morador individual em domicílio coletivo (por muitas vezes no texto nos referiremos a essas duas categorias apenas como responsável por domicílio). Da não-resposta a essa questão, podem surgir diversos efeitos sobre análises a serem feitas. Daí, surge a necessidade de que seja feita a imputação de valores de rendimento dos não-respondentes.

Dada a magnitude de uma pesquisa como o Censo Demográfico 2000, fez-se necessária a adoção de uma metodologia de imputação que permitisse o processamento rápido e automatizado da grande massa de dados existente, além de atingir o objetivo principal de corrigir os possíveis efeitos causados pela não-resposta. Com essa finalidade, foi desenvolvida uma metodologia baseada na técnica de Árvores de Regressão (BREIMAN et al., c1984).

Aspectos gerais

A não-resposta é um dos mais comuns erros entre os não-amostrais de uma pesquisa, sendo bem freqüente em países mais desenvolvidos, embora ultimamente venha crescendo em países como o Brasil. Em geral, perguntas sobre rendimentos são mais sujeitas à não-resposta do que as demais perguntas existentes em uma pesquisa como o censo demográfico.

A não-resposta pode ser de dois tipos: completa, quando o total das informações a serem obtidas de uma unidade de pesquisa não é coletada; parcial, quando apenas uma parte das informações não é coletada. No caso do Conjunto Universo do Censo Demográfico 2000, a não-resposta do rendimento do responsável pelo domicílio pode ser vista como parcial, pois as não-respostas que tenham ocorrido nas demais perguntas do questionário foram imputadas com o uso do sistema DIA antes da etapa de imputação de rendimento. A não-utilização do DIA para a imputação de rendimento, justifica-se pelo fato deste ser um sistema adequado à imputação de variáveis categóricas ou numéricas discretas.

Uma pergunta fundamental, cuja resposta justifica em muito a execução de um procedimento de imputação de rendimento, é: quais os efeitos de ignorar-se a não-resposta ao se fazer inferências ou análises sobre o rendimento dos responsáveis por domicílios? No caso da não-resposta não diferencial, isto é, ao acaso, o seu efeito é o de aumento de variância das estimativas obtidas para parâmetros populacionais. No caso da não-resposta diferencial, o seu impacto se dá sob a forma de vício nas estimativas obtidas, com esse vício crescendo com a taxa de não resposta e com a diferença entre respondentes e não-respondentes.

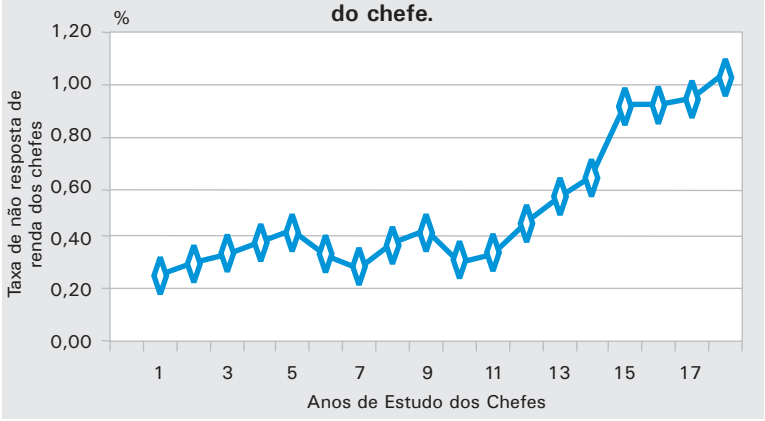
Análises realizadas com dados do Censo Demográfico 1991, mostraram que a não-resposta nos rendimentos dos chefes de domicílio se dava de forma diferencial em relação a algumas das variáveis presentes no questionário. O Gráfico 12.2 mostra que a taxa de não-resposta de rendimento dos chefes de domicílio no Censo de

1991, cresce conforme aumentam os anos de estudo dos chefes. Em geral, foi possível verificar com os dados do Censo de 1991 que as taxas de não-resposta mais elevadas estavam associadas a valores de variáveis que caracterizavam níveis mais altos de rendimento. Donde concluiu-se que quanto maior o rendimento dos chefes de domicílio maior era a probabilidade de não-resposta do rendimento.

São duas as alternativas para lidar com o problema da não-resposta diferencial: uso de estimadores adequados para dados faltantes (LITTLE; RUBIN, 1987) e uso de métodos de imputação (substituição de valores estimados em cada caso individual). No caso de não-resposta parcial de um item/variável, a preferência das agências de estatísticas oficiais é geralmente por métodos baseados em imputação. Tal preferência se deve à maior simplicidade dessa alternativa no processamento posterior dos dados, particularmente quando estes precisam ser publicados na forma de arquivos de microdados com as informações de cada pessoa, individualmente. Albieri (1992) investigou a aplicação de vários métodos para imputação do rendimento na Pesquisa Mensal de Emprego do IBGE.

O método de imputação adotado trabalha com a idéia de estabelecer uma relação entre rendimentos declarados dos responsáveis por domicílios e um grupo de variáveis do conjunto universo cujos valores são conhecidos para todos os domicílios pesquisados, e a partir dessa relação imputar valores de rendimento para os não-respondentes.

Gráfico 12.2 - Taxa de não-resposta de rendimento dos chefes de domicílio no Censo Demográfico 1991, segundo os anos de estudo do chefe.



Dentre as variáveis existentes para o Conjunto Universo algumas, foram selecionadas como possíveis variáveis explicativas do rendimento dos responsáveis por domicílios, sendo este conjunto diverso o suficiente para descrever de forma satisfatória as diferentes relações com o rendimento existentes ao longo do país. A seguir, é apresentado o conjunto dessas variáveis selecionadas para utilização na processo de imputação, com a descrição de cada variável precedida da respectiva nomenclatura adotada para ela durante :

1. IDADANO – Idade em anos do responsável pelo domicílio;

2. ANOSEST – Anos de estudo do responsável pelo domicílio;

3. SEXO – Sexo do responsável pelo domicílio;

4. ESPECIE – Espécie de domicílio: particular permanente; particular improvisado; coletivo;

5. TIPODOM – Tipo de domicílio: casa; apartamento; cômodo;

6. TOTPESDO – Total de moradores no domicílio;

7. EMPREDOM – Total de empregados domésticos residentes no domicílio;

8. CONDDOM – Condição do domicílio: próprio - já pago; próprio – ainda pagando; alugado; cedido por empregador; cedido de outra forma; outra condição;

9. QTDBANH – Quantidade de banheiros existentes no domicílio;

10. SANITAR – Indicadora de existência de sanitário no domicílio com zero banheiro;

11. ABASTEC – Tipo de abastecimento de água: rede geral; poço ou nascente (na propriedade); outra;

12. TIPOCAN – Tipo de canalização de água: canalizada em pelo menos um cômodo; canalizada só na propriedade ou terreno; não canalizada;

13. TIPOESC – Tipo de ligação do escoadouro do banheiro ou sanitário do domicílio: rede geral de esgoto ou pluvial; fossa séptica; fossa rudimentar; vala; rio, lago ou mar; outro escoadouro;

14. LIXO – Tipo de coleta de lixo do domicílio: coletado por serviço de limpeza; colocado em caçamba de serviço de limpeza; queimado (na propriedade); enterrado (na propriedade); jogado em terreno baldio ou logradouro; jogado em rio, lago ou mar; tem outro destino;

15. TIPOSET – Tipo de setor censitário em que se situa o domicílio.

Pode-se notar que parte das variáveis utilizadas refere-se diretamente ao responsável pelo domicílio (de 1 a 3); outra parte é referente a características do domicílio (de 4 a 10); enquanto as demais referem-se a local onde se situa o domicílio (de 11 a 15).

Metodologia

Como já foi dito anteriormente, a metodologia empregada baseia-se na técnica de Árvores de Regressão. A seguir, é dada uma breve idéia a respeito do funcionamento da técnica.

Considere-se a seguinte situação: é preciso predizer o rendimento de uma pessoa. Se for levado em conta apenas o fato de que essa pessoa viva e trabalhe no município do Rio de Janeiro, um preditor bastante "grosseiro" seria o rendimento médio da população desse município. O problema com esse preditor é que ele teria uma precisão muito pequena, ou seja, uma dispersão muito grande, isso se comparado a outros possíveis preditores que levassem em conta variáveis explicativas do rendimento das pessoas residentes no município do Rio de Janeiro.

Portanto, para melhorar a predição pode-se usar outras informações sobre a pessoa. Por exemplo, se for considerado não só o lugar onde ela reside, mas também informações como: idade; nível de instrução; sexo; etc., pode-se assim melhorar a qualidade do preditor. Novamente seria calculada uma média de rendimentos, mas agora sobre uma população bem mais restrita e homogênea.

Esta é a idéia básica das técnicas de regressão: calcular médias em subgrupos (estratos) definidos por variáveis explicativas (covariáveis), obtendo um preditor mais preciso da variável resposta do que o obtido caso não fossem usadas informações sobre essas covariáveis.

Continuando com o exemplo, que perguntas deveríam ser feitas a fim de melhor predizer o rendimento da pessoa? Ou seja, que variáveis explicativas escolher e como fazer a pergunta? Deve ser lembrado que não é permitido perguntar diretamente sobre o rendimento.

Suponha-se que só seja possível perguntar sobre um conjunto dado de variáveis explicativas. Mais ainda, as perguntas são específicas e só podem ser do seguinte tipo:

- no caso de variável numérica: se está abaixo de um valor escolhido (por exemplo: idade ≤ 27 anos; anos de estudo ≤ 7 anos, etc); e
- no caso de variável categórica: se pertence a um subconjunto de categorias.

A resposta a cada pergunta formulada será sim ou não. Mas como escolher as perguntas a fazer? Escolher uma pergunta implica duas escolhas: a da variável explicativa e a de como formular a pergunta a respeito da variável selecionada. Note que no caso de predições de rendimento para indivíduos, as respostas às perguntas definem, passo a passo, estratos cada vez menores de indivíduos. Portanto, para o objetivo é importante escolher as perguntas de modo que esses estratos sejam cada vez mais homogêneos em relação ao rendimento.

Na técnica de Árvores de Regressão, para estabelecer a melhor seqüência de perguntas definidoras dos estratos, parte-se de uma amostra onde sejam conhecidas para cada indivíduo o seu rendimento e os valores das variáveis explicativas. Essa amostra recebe o nome de amostra de treinamento, pois a partir dela é que "se entende" a relação entre o rendimento e as covariáveis adotadas.

Suponha-se, ainda, que seja conhecido um critério numérico D para comparar partições de grupos em dois subgrupos e que o valor de D só dependa dos valores dos rendimentos nos subgrupos definidos. Então, usando a amostra de treinamento, pode-se usar o seguinte procedimento:

- na amostra de treinamento, faz-se todas as perguntas possíveis sobre cada uma das variáveis explicativas, obedecendo as especificações acima definidas para as perguntas. Para cada partição definida por cada pergunta calcula-se o valor do critério. Escolhe-se a pergunta que minimize D. Observe que só é preciso um número finito de perguntas, pois os subgrupos definidos só seriam modificados quando um indivíduo mudasse de grupo, o que ocorreria quando o "ponto de corte" coincidisse com um valor da variável na amostra de treinamento;
- o mesmo procedimento acima seria aplicado em cada um dos dois subgrupos obtidos, sendo sucessivamente geradas partições binárias no grupo de indivíduos da amostra de treinamento. Note que nesse segundo passo, a cada definição de partição é necessário escolher em qual subgrupo particionar. Para isso, bastaria calcular os valores de D referentes às possíveis partições e selecionar aquela para a qual fosse minimizado o valor do critério; e

- por último, é preciso definir um critério de parada para o processo de partições. Possibilidades: limite inferior para o contingente nos subgrupos; ou o fato de que uma nova partição traga "pouca melhora" em termos do critério adotado.

O procedimento acima descrito é uma síntese do funcionamento da técnica de Árvores de Regressão, podendo ser representado por uma árvore binária. Na figura 12.2 é exemplificada uma árvore de regressão, onde para uma amostra de treinamento fictícia (ver tabela 12.7, a seguir) os seus componentes têm o rendimento explicado pelas seguintes covariáveis: sexo; idade; anos de estudo.

Tabela 12.7 - Informações individuais de uma amostra de treinamento fictícia

Sexo	Renda	Idade	Anos de estudo
Masculino	100,00	18	4
Masculino	200,00	20	8
Masculino	200,00	24	6
Masculino	150,00	25	4
Masculino	450,00	30	11
Masculino	300,00	32	8
Masculino	200,00	35	1
Masculino	200,00	46	4
Masculino	1 200,00	63	11
Feminino	200,00	17	8
Feminino	50,00	22	1
Feminino	80,00	25	-
Feminino	150,00	32	4
Feminino	200,00	33	8
Feminino	400,00	35	11
Feminino	300,00	39	8
Feminino	280,00	44	4
Feminino	280,00	49	8
Feminino	120,00	52	5
Feminino	100,00	71	4

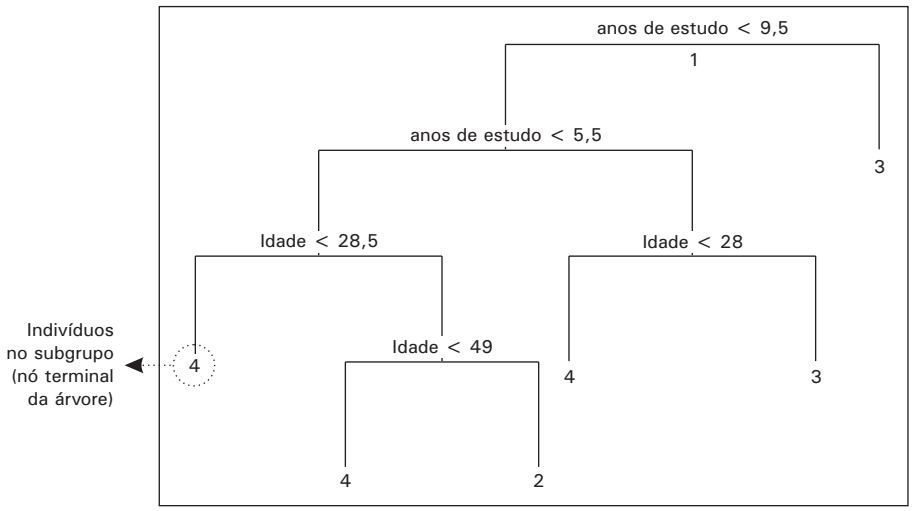
Fonte: IBGE, Diretoria de Pesquisas, Departamento de Metodologia.

Os principais aspectos a comentar sobre a figura 12.2 são:

- a primeira partição dá-se pela separação dos indivíduos com menos de 9,5 anos de estudo dos com mais de 9,5 anos de estudo;
- a segunda partição é feita dentro do grupo de indivíduos com menos de 9,5 anos de estudo, separando-se os que têm menos de 5,5 anos de estudo dos demais indivíduos do estrato;

- as partições são feitas sucessivamente até a condição de parada ser atingida, o que para este exemplo foi estabelecida como sendo a existência de um mínimo de dois indivíduos por nó terminal da árvore; e
- nota-se que a variável "sexo" não é utilizada para a construção da árvore. Isso ilustra o fato de que, na técnica de Árvores de Regressão, não necessariamente todas as covariáveis presentes na amostra de treinamento devam ser utilizadas. A técnica tem, por si só, a capacidade de selecionar as variáveis explicativas mais "poderosas" no sentido de explicar a variável resposta.

Figura 12.2 - Exemplo de Árvore de Regressão



O critério D adotado foi a deviance (soma de desvios quadráticos), que pode ser assim definida:

$$D = \sum_{j} \sum_{i} (y_{ij} - \bar{y}_{j})^2$$

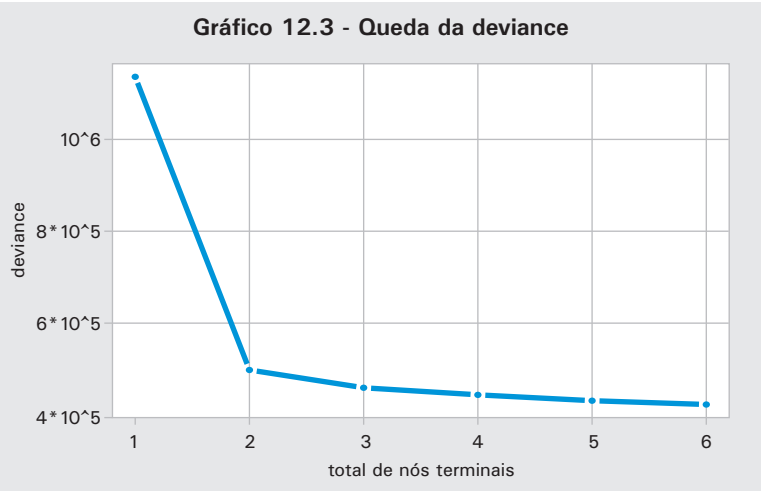
onde:

- y_{ij} é o valor da variável resposta observado para o indivíduo i pertencente ao estrato (nó terminal) j;
- $\bar{y}_{j}$ é a média da variável resposta no estrato j.

No Gráfico 12.3 é apresentado o comportamento da *deviance* com o aumento do número de partições na árvore de regressão construída para a amostra de treinamento adotada como exemplo. Conforme pode-se notar, há uma queda acentuada no valor de D ao particionar-se o grupo inicial em dois estratos,

com as partições seguintes trazendo "melhoras" cada vez menores na obtenção de estratos mais homogêneos em relação ao rendimento. Esses gráficos podem ser utilizados na escolha do número de nós terminais a ser adotado.

Gráfico 12.3 - Queda da deviance



Conhecidas as perguntas a fazer, como então predizer um rendimento desconhecido, isto é, como imputar os rendimentos dos não respondentes?

Dado que são conhecidos os valores das variáveis explicativas para os que não declararam seus valores de rendimentos, uma solução seria localizar essas pessoas nos nós terminais e, em seguida, imputar o rendimento de cada uma pelo rendimento médio em seu respectivo "nó". Porém, tal solução possui o inconveniente de não manter a distribuição original da variável resposta em cada estrato, visto que seria sempre imputado o rendimento médio no estrato. Por esse motivo adota-se o procedimento denominado hot-deck aleatório, onde para cada indivíduo não-respondente seleciona-se aleatoriamente um "doador de rendimento" dentro de seu nó terminal, e imputa-se seu rendimento pelo rendimento do doador.

Aplicação e conclusões

Para o processo de imputação de rendimento dos responsáveis, utilizaram-se os mesmos lotes de registros definidos para a crítica e imputação de dados do Censo Demográfico 2000, realizadas com o uso do sistema DIA. Esses lotes correspondem a uma partição do conjunto universo de respondentes, obedecendo os domínios das Unidades da Federação (UF), isto é, um mesmo lote não contém registros de mais de uma UF. Para o processo de imputação de rendimento foram utilizados somente os registros de responsáveis por domicílio ou morador individual em domicílio coletivo, além de terem sido excluídos de cada lote os registros cujos rendimentos estavam fora das cercas construídas para detectar outliers (valores atípicos). Nos 526 lotes utilizados na imputação de rendimento havia o total de 45.280.240 registros, com o menor lote possuindo 10.094 registros e o maior 103.248 registros. A distribuição da quantidade de lotes por Unidades da Federação pode ser vista na tabela 12.8.

Tabela 12.8 - Número de lotes para imputação, segundo as Unidades da Federação

Unidades da Federação	Lotes	Unidades da Federação	Lotes
Brasil	**526**	Alagoas	9
Rondônia	5	Sergipe	6
Acre	3	Bahia	36
Amazonas	8	Minas Gerais	54
Roraima	2	Espírito Santo	10
Pará	16	Rio de Janeiro	48
Amapá	2	São Paulo	117
Tocantins	4	Paraná	31
Maranhão	15	Santa Catarina	18
Piauí	8	Rio Grande do Sul	34
Ceará	21	Mato Grosso do Sul	7
Rio Grande do Norte	8	Mato Grosso	8
Paraíba	10	Goiás	16
Pernambuco	23	Distrito Federal	7

Fonte: IBGE, Censo Demográfico 2000.

Para cada um dos 526 lotes foi aplicado o procedimento de imputação baseado em árvores de regressão, descrito acima. Esse procedimento foi implementado com o software *S-Plus* e executado em ambiente operacional Windows 98. Como os lotes de registros residiam em arquivos do ambiente operacional OS/390 (mainframe IBM), foi desenvolvida, utilizando o software SAS e seus recursos para a conexão desses dois ambientes operacionais, uma rotina computacional para automatizar todo o processo de produção dessa imputação, constituído das seguintes etapas: a) preparação do arquivo de entrada para o *S-Plus*; b) ativação do *S-Plus* para a imputação propriamente dita; c) transferência dos resultados para o ambiente OS/390 e d) atualização dos registros nos lotes originais com os valores imputados.

A regra de parada na construção das árvores de regressão baseou-se no número máximo de nós terminais permitido nas árvores e no contingente populacional mínimo exigido em cada nó terminal. Visto que seria impraticável a análise dos gráficos de queda da *deviance* para cada uma das 526 árvores, uma das regras de parada adotada foi a da partição de cada lote em no máximo 25 estratos. Análises preliminares com dados do Censo de 1991 indicaram ser este um número de nós terminais para o qual, em geral, não haveria "ganhos consideráveis" com novas partições. Quanto aos contigentes populacionais de cada estrato, foi estipulado que estes deveriam ser de, no mínimo, 100 pessoas.

Tabela 12.9 - Estatísticas descritivas das taxas de não-resposta nos lotes de imputação de rendimento dos responsáveis por domicílios

Taxas de não-resposta nos lotes de imputação (%)					
Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo
0,54	1,17	1,60	1,78	2,15	6,24

Fonte: IBGE, Censo Demográfico 2000.

Além das estatísticas referentes às taxas de não-resposta nos lotes, apresentadas na Tabela 12.9, há que se dizer ainda que a taxa geral de não-resposta de rendimento entre os responsáveis por domicílios foi de 1,75%.

O Gráfico 12.4 apresenta os resultados da participação de cada variável explicativa nas árvores de regressão construídas, com os dados utilizados na construção do gráfico podendo ser vistos na Tabela 12.10. Essa participação é vista sob o ponto de vista da ordem em que a variável gerou uma partição na árvore pela primeira vez. Por exemplo, as variáveis QTDBANH e ANOSEST geraram a 1ª ou a 2ª partição em 94,7% e 87,6% das árvores, respectivamente.

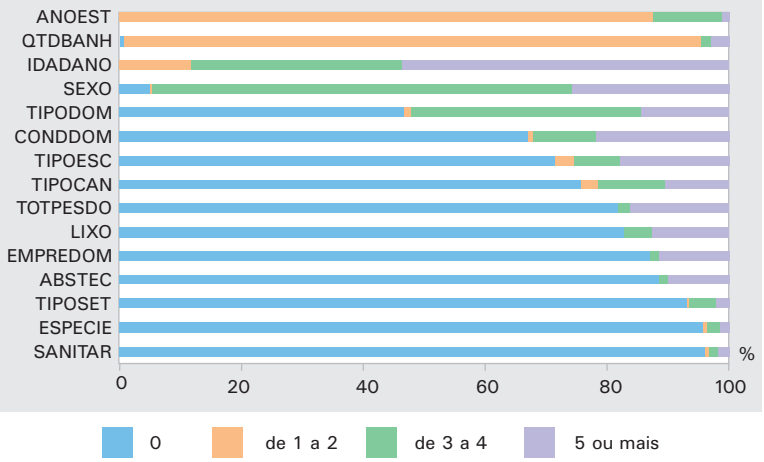
Tabela 12.10 - Resumo da participação das variáveis nas árvores de regressão

Variável	Ordem de entrada na variável árvore (%)				
	1	2	3 ou 4	5 ou mais	Não entrou
QTDBANH	53,80	40,87	1,71	2,85	0,76
ANOSEST	39,92	47,72	11,41	0,95	0,00
IDADANO	3,99	3,23	34,79	53,42	4,56
TIPOESC	1,33	2,09	7,41	17,87	71,29
TIPOCAN	0,57	2,28	11,03	10,46	75,67
TIPODOM	0,19	0,76	37,64	14,45	46,96
SANITAR	0,19	0,76	1,33	1,71	96,01
ESPECIE	0,00	0,76	2,09	1,52	95,63
SEXO	0,00	0,57	68,63	25,86	4,94
CONDDOM	0,00	0,38	10,46	21,86	67,30
TIPOSET	0,00	0,38	4,18	2,09	93,35
TOTPESDO	0,00	0,19	1,90	16,16	81,75
LIXO	0,00	0,00	4,75	12,36	82,89
ABASTEC	0,00	0,00	1,52	10,08	88,40
EMPREDOM	0,00	0,00	1,14	11,60	87,26

Fonte: IBGE, Diretoria de Pesquisas, Departamento de Metodologia.

Nota: As variáveis V0407, V0408 e V4093 só foram investigadas para a pessoa responsável pelo domicílio ou individual em domicílio coletivo.

Gráfico 12.4 - Ordem de entrada das variáveis



Ainda a respeito das variáveis ANOSEST e QTDBANH, nota-se que a primeira foi selecionada na construção das árvores de todos os lotes, enquanto a segunda não entrou em menos de 1% das árvores construídas. As variáveis IDADANO e SEXO entraram em mais de 95% das árvores, em geral foram a 3ª ou 4ª a ser selecionada, porém a variável SEXO foi com maior freqüência mais importante do que IDADANO. Por outro lado, as variáveis TOTPESDO; LIXO; EMPREDOM; ABASTEC; TIPOSET; ESPECIE e SANITAR, não foram incluídas nas árvores de mais de 80% dos lotes.

Conforme pôde-se notar pelos resultados apresentados, as variáveis diretamente relacionadas à pessoa do responsável pelo domicílio foram as que apresentaram maior poder de predição do rendimento desta pessoa, excetuando-se aí, é claro, a variável QTDBANH.

Como forma de avaliar a qualidade do resultado da imputação em cada lote foi aplicado o teste estatístico de Kolmogorov-Smirnov (LEHMANN, c1975). Este teste visa verificar se duas amostras de dados provêm de uma mesma população. No nosso caso, como se tinha o objetivo de não alterar a distribuição do rendimento em cada subgrupo formado, então o teste foi aplicado para se comparar os vetores de rendimentos em cada subgrupo antes e depois da execução do procedimento de imputação. Cada lote só teve seu respectivo processo de imputação aprovado se o teste de Kolmogorov-Smirnov indicasse que os rendimentos antes e depois da imputação, em cada estrato formado, apresentavam a "mesma distribuição".

Outro ponto a ser colocado diz respeito ao tratamento dado à imputação de rendimentos nulos. Visto que seria possível a categorização da variável de rendimentos em nulos e não-nulos, haveria a possibilidade do sistema DIA ser utilizado para a imputação de valores nulos de rendimentos. Porém, essa solução não foi adotada porque avaliou-se que, no caso do conjunto universo, o conjunto de variáveis explicativas disponíveis não possuía boa capacidade de predição da variável de rendimento dicotomizada (sem rendimento ou com rendimento positivo). Posto isso, optou-se por realizar a imputação de rendimentos nulos dentro do procedimento de imputação estabelecido.

Finalizando, são apresentadas a seguir algumas estatísticas referentes às distribuições nos lotes do percentual de responsáveis contido em cada estrato (tabela 12.11) e às distribuições das taxas de imputação nos estratos obtidos nas árvores de regressão construídas (tabela 12.12). Um aspecto a comentar no Tabela 12.12 diz respeito ao fato de ter sido observado extrato com 100% de não-resposta, o que acarretou na inexistência de "doadores de rendimento". Tal problema foi solucionado com a seleção de doador no grupo de responsáveis do qual foi gerado o estrato.

Tabela 12.11 - Estatísticas descritivas dos percentuais de responsáveis nos estratos em relação ao tamanho dos lotes de imputação de rendimento dos responsáveis por domicílios

Percentual de responsáveis no estrato em relação ao tamanho do lote (%)					
Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo
0,001	0,580	1,530	3,990	4,220	61,120

Fonte: IBGE, Censo Demográfico 2000.

Tabela 12.12 - Estatísticas descritivas das taxas de não-resposta nos estratos das árvores de regressão dos lotes de imputação de rendimento dos responsáveis por domicílios

Percentual de responsáveis no estrato em relação ao tamanho do lote					
Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo
0,00	0,67	1,15	1,54	1,89	100,00

Fonte: IBGE, Censo Demográfico 2000.

12.4 Resultados preliminares da amostra

Os resultados preliminares da amostra do Censo Demográfico 2000 foram divulgados em maio de 2002, correspondendo à Tabulação Avançada e à Fecundidade e Mortalidade Infantil.

Diferentemente da Fecundidade e Mortalidade Infantil, onde foram utilizados para a obtenção das estimativas os dados de toda a amostra do censo, na Tabulação Avançada recorreu-se à elaboração de uma subamostra.

Com exceção dos dados referentes à estrutura familiar e sexo, os demais, utilizados para a obtenção desses resultados, ainda não tinham sido submetidos ao processo de crítica e imputação.

12.4.1 Tabulação Avançada

O objetivo da Tabulação Avançada foi fornecer, antecipadamente à divulgação dos resultados da amostra do censo, estimativas para um conjunto de tabelas com variáveis do questionário da amostra para o total do País, Grandes Regiões e Unidades da Federação. Para isso, foi retirada uma subamostra da amostra do Censo Demográfico 2000, constituída por uma amostra de setores censitários, com os respectivos domicílios e pessoas neles pesquisados, que preencheram o questionário da amostra, cujas informações ainda não haviam sido submetidas a todos os processos de crítica eletrônica.

Portanto, a Tabulação Avançada contém resultados preliminares da amostra do Censo Demográfico 2000, que estavam sujeitos a alterações quando da versão definitiva.

Plano amostral da Tabulação Avançada

O plano amostral da Tabulação Avançada consistiu em uma amostragem estratificada simples de setores censitários em cada unidade da federação (cada um dos 26 estados e o Distrito Federal). Em cada Unidade da Federação foram definidos até três estratos de acordo com a situação e o tipo do setor, a saber: setores rurais, setores urbanos não-especiais, e setores urbanos de aglomerados subnormais, quando existentes. Dentro de cada estrato, os setores foram selecionados por amostragem aleatória simples. A seleção dos setores foi feita utilizando-se o algoritmo sugerido por Fan, Muller e Rezucha em 1962, que está descrito em Särndal, Swensson e Wretman (c1992, p. 66).

O tamanho da amostra de setores em cada Unidade da Federação (UF) foi definido com base em estudos descritos em Albieri, Martelotte e Duarte (2000) e em Albieri (1999b). Nestes estudos ficou definido que, para a estimação de características para Unidades da Federação (UF) com precisão razoável, o número de

setores a ser utilizado seria o equivalente ao da amostra da Pesquisa Nacional por Amostra de Domicílios – PNAD, e que o tamanho mínimo da amostra de uma UF seria de 50 setores. A alocação da amostra de setores nos três estratos foi proporcional ao número de setores existentes em cada estrato na população, sendo considerado um mínimo de dois setores, para permitir a estimação do erro amostral.

A Tabela 12.13 a seguir mostra o número de setores do Censo Demográfico e da Tabulação Avançada por Unidade da Federação.

Tabela 12.13 - Número de setores do Censo Demográfico 2000 e da Tabulação Avançada, segundo as Unidades da Federação

Unidades da Federação	Setores Do Censo 2000 (1)	Setores da Tabulação Avançada (1)				Fração amostral de setores (%)
		Total	Estratos			
			Rural	Urbano	Aglomerado subnormal	
Brasil	**214 319**	**4 359**	**1 312**	**2 897**	**150**	**2**
Rondônia	1 888	57	28	29	-	3
Acre (1)	552	51	21	30	-	9,2
Amazonas	3 236	104	33	64	7	3,2
Roraima	476	51	20	31	-	10,7
Pará	6 083	121	50	60	11	2
Amapá	452	51	13	35	3	11,3
Tocantins	1 364	58	26	32	-	4,3
Maranhão	6 398	114	65	46	3	1,8
Piauí	3 708	97	51	42	4	2,6
Ceará	7 947	193	65	117	11	2,4
Rio Grande do Norte	2 633	91	35	53	3	3,5
Paraíba	4 162	114	48	63	3	2,7
Pernambuco	8 541	233	85	142	6	2,7
Alagoas	2 600	94	38	53	3	3,6
Sergipe	2 220	87	33	51	3	3,9
Bahia	15 315	277	118	154	5	1,8
Minas Gerais	22 469	352	100	242	10	1,6
Espírito Santo	3 196	109	31	75	3	3,4
Rio de Janeiro	20 607	293	19	247	27	1,4
São Paulo	49 303	423	47	354	22	0,9
Paraná	13 005	212	68	139	5	1,6
Santa Catarina	6 794	173	52	118	3	2,5
Rio Grande do Sul	16 837	316	94	216	6	1,9
Mato Grosso do Sul	2 710	120	37	80	3	4,4
Mato Grosso	3 309	136	51	82	3	4,1
Goiás	5 960	263	75	185	3	4,4
Distrito Federal	2 554	169	9	157	3	6,6

Fonte: IBGE, Censo Demográfico 2000.

(1) Setores com pelo menos uma pessoa recenseada no Censo Demográfico 2000.

Em cada setor todos os domicílios e pessoas nele pesquisados, através do Questionário da Amostra, foram processados e incluídos na amostra da Tabulação Avançada.

Nas Tabelas 12.14 e 12.15 a seguir, apresenta-se o número de domicílios e de pessoas do Censo Demográfico e da Tabulação Avançada por Unidade da Federação.

Tabela 12.14 - Número de domicílios do Censo Demográfico 2000 e da Tabulação Avançada, segundo as Unidades da Federação

Unidades da Federação	Domicílios do Censo 2000 (1)	Domicílios da Tabulação Avançada (1)				Fração amostral de domicílios (%)
		Total	Estratos			
			Rural	Urbano	Aglomerado subnormal	
Brasil	**45 507 516**	**108 989**	**25 238**	**80 339**	**3 412**	**0,24**
Rondônia	354 391	1 149	448	701	-	0,32
Acre	131 580	1 539	575	964	-	1,17
Amazonas	580 900	2 092	556	1 383	153	0,36
Roraima	76 681	1 100	461	639	-	1,43
Pará	1 332 248	2 810	1 011	1 547	252	0,21
Amapá	100 765	1 203	137	916	150	1,19
Tocantins	285 701	1 710	377	1 333	-	0,60
Maranhão	1 246 715	2 463	960	1 430	73	0,20
Piauí	665 808	2 559	987	1 466	106	0,38
Ceará	1 773 393	4 866	1 510	3 127	229	0,27
Rio Grande do Norte	678 652	3 029	922	2 031	76	0,45
Paraíba	857 989	3 052	1 000	1 974	78	0,36
Pernambuco	1 994 041	5 946	1 496	4 268	182	0,30
Alagoas	658 873	2 805	870	1 902	33	0,43
Sergipe	442 256	2 138	675	1 394	69	0,48
Bahia	3 214 292	6 902	2 358	4 404	140	0,21
Minas Gerais	4 837 296	9 881	2 055	7 592	234	0,20
Espírito Santo	851 014	3 342	890	2 356	96	0,39
Rio de Janeiro	4 315 737	5 993	300	5 210	483	0,14
São Paulo	10 564 745	9 660	968	8 274	418	0,09
Paraná	2 709 523	5 516	1 218	4 221	77	0,20
Santa Catarina	1 518 651	5 028	1 270	3 714	44	0,33
Rio Grande do Sul	3 091 643	6 751	1 502	5 132	117	0,22
Mato Grosso do Sul	577 362	3 056	612	2 397	47	0,53
Mato Grosso	669 676	3 472	746	2 517	209	0,52
Goiás	1 420 822	7 286	1 186	6 010	90	0,51
Distrito Federal	556 762	3 641	148	3 437	56	0,65

Fonte: IBGE, Censo Demográfico 2000.

(1) Número de domicílios com pelo menos uma pessoa recenseada no Censo Demográfico 2000.

Tabela 12.15 - Número de pessoas do Censo Demográfico 2000 e da Tabulação Avançada, segundo as Unidades da Federação

(continua)

Unidades da Federação	Pessoas do Censo 2000	Pessoas da Tabulação Avançada				Fração amostral de pessoas (%)
		Total	Estratos			
			Rural	Urbano	Aglomerado subnormal	
Brasil	**169 799 170**	**423 049**	**108 938**	**300 268**	**13 843**	**0,25**
Rondônia	1 379 787	4 450	1 800	2 650	-	0,32
Acre	557 526	6 295	2 478	3 817	-	1,13
Amazonas	2 812 557	9 940	3 087	6 164	689	0,35
Roraima	324 397	4 534	1 908	2 626	-	1,40
Pará	6 192 307	13 395	5 141	7 185	1 069	0,22
Amapá	477 032	5 802	673	4 375	754	1,22
Tocantins	1 157 098	6 717	1 474	5 243	-	0,58
Maranhão	5 651 475	11 644	4 642	6 672	330	0,21
Piauí	2 843 278	10 939	4 339	6 141	459	0,38
Ceará	7 430 661	20 843	7 145	12 763	935	0,28
Rio Grande do Norte	2 776 782	12 711	4 252	8 127	332	0,46

Tabela 12.15 - Número de pessoas do Censo Demográfico 2000 e da Tabulação Avançada, segundo as Unidades da Federação

(conclusão)

Unidades da Federação	Pessoas do Censo 2000	Pessoas da Tabulação Avançada				Fração amostral de pessoas (%)
		Total	Estratos			
			Rural	Urbano	Aglomerado subnormal	
Paraíba	3 443 825	12 676	4 598	7 776	302	0,37
Pernambuco	7 918 344	24 575	7 283	16 508	784	0,31
Alagoas	2 822 621	12 746	4 528	8 071	147	0,45
Sergipe	1 784 475	9 231	3 227	5 697	307	0,52
Bahia	13 070 250	29 078	10 821	17 666	591	0,22
Minas Gerais	17 891 494	37 835	8 887	28 026	922	0,21
Espírito Santo	3 097 232	12 551	3 700	8 467	384	0,41
Rio de Janeiro	14 391 282	20 633	1 196	17 686	1 751	0,14
São Paulo	37 032 403	33 718	3 191	28 793	1 734	0,09
Paraná	9 563 458	19 439	4 451	14 701	287	0,20
Santa Catarina	5 356 360	18 151	5 169	12 763	219	0,34
Rio Grande do Sul	10 187 798	22 330	5 273	16 635	422	0,22
Mato Grosso do Sul	2 078 001	10 990	2 237	8 543	210	0,53
Mato Grosso	2 504 353	12 871	2 683	9 506	682	0,51
Goiás	5 003 228	25 749	4 243	21 203	303	0,51
Distrito Federal	2 051 146	13 206	512	12 464	230	0,64

Fonte: IBGE, Censo Demográfico 2000.

Expansão da amostra e cálculo dos pesos amostrais da Tabulação Avançada

Numa pesquisa por amostra o que se busca são estimativas dos valores de determinados parâmetros populacionais de interesse, por meio da investigação de apenas uma parte das unidades dessa população. As tabelas divulgadas na Tabulação Avançada são formadas pelo cruzamento de variáveis relativas a pessoas, domicílios e famílias, classificadas segundo categorias indicadoras de faixa etária, sexo, religião, rendimento, localização, etc. Cada uma das células das tabelas teve seu valor estimado, juntamente com uma medida de precisão de estimativa dada pelo coeficiente de variação.

O processo de cálculo das estimativas é também conhecido como expansão da amostra e depende da determinação dos pesos associados a cada unidade amostrada. Os pesos usuais dados pelo plano amostral, definidos como o inverso das probabilidades de inclusão de cada unidade da população na amostra, são os pesos mais simples, que podem ser utilizados para a expansão dos resultados de uma pesquisa por amostra. No caso do desenho amostral da Tabulação Avançada, estes pesos podem ser escritos como:

$$d_{hij} = \frac{1}{\pi_{hij}} = \frac{M_h}{m_h}\frac{N_{hi}}{n_{hi}} \quad ,$$

onde:

d_{hij} é o peso do domicílio j, do setor i, do estrato h na amostra;

π_{hij} é a probabilidade de inclusão do domicílio j, do setor i, do estrato h na amostra;

M_h e m_h são, respectivamente, o número de setores no estrato h no Censo Demográfico 2000 e na amostra da Tabulação Avançada;

N_{hi} e n_{hi} são, respectivamente, o número de domicílios no setor i do estrato h no Censo Demográfico 2000 e na amostra.

Em cada unidade da federação existem, no máximo, três estratos[3]. Esses pesos utilizam apenas informações do plano amostral.

Na Tabulação Avançada os pesos utilizados foram os pesos definidos como descrito, calibrados de modo a se ajustarem às informações auxiliares fornecidas pela investigação censitária, realizada pela aplicação das perguntas comuns aos dois tipos de questionário, que são feitas para todos os domicílios da população. Foram utilizados como variáveis de calibração os totais de homens e mulheres e o total de domicílios por estrato, já que tais informações estavam disponíveis para todos os domicílios da população e, conseqüentemente, da amostra.

Os pesos amostrais indicam quantas unidades da população cada unidade da amostra representa. A calibração dos pesos foi feita de maneira que os pesos calibrados ficassem o mais próximo possível dos pesos dados pelo desenho amostral, impondo-se como limite mínimo o valor 1, ou seja, nestes casos cada unidade representa apenas ela mesma.

A razão para se optar pelo uso de pesos calibrados vem do fato de estes produzirem estimativas mais precisas e mais consistentes com os valores conhecidos do Censo Demográfico 2000. Os totais, para os estratos, das variáveis usadas na calibração, quando estimados pelos pesos calibrados, coincidem com os valores conhecidos para toda a população.

O método de calibração usado é um processo com restrições não-lineares, nas variáveis de calibração com limites nos valores dos pesos calibrados. Detalhes podem ser vistos na publicação *Generalised estimation system* (1998), do Statistics Canada.

Os pesos foram calculados no nível do domicílio, sendo atribuídos também a cada um de seus moradores.

Cálculo das estimativas pontuais da Tabulação Avançada

As estimativas de totais para as células das tabelas da Tabulação Avançada foram calculadas por:

$$\hat{Y} = \sum_{h=1}^{H} \sum_{i=1}^{m_h} \sum_{j=1}^{n_{hi}} d^{*}_{hij}\ y_{hij}$$

onde:

$d^{*}_{hij} = peso\ \ calibrado\ \ para\ a\ \ unidade\ \ j\ \ do\ setor\ \ i\ do\ estrato\ \ h$;

$H = número\ \ de\ \ estratos$; e

a) Para variáveis categóricas:

$$y_{hij} = \begin{cases} 1,\ se\ a\ unidade\ \ j\ do\ setor\ i\ do\ estrato\ h\ \ pertence\ à\ categoria\ em\ questão \\ 0,\ se\ a\ unidade\ \ j\ do\ setor\ i\ do\ estrato\ h\ não\ \ pertence\ à\ categoria\ em\ questão \end{cases}$$

[3] Em algumas Unidades da Federação, como Rondônia, Acre, Roraima e Tocantins, não havia setor, com pelo menos uma pessoa recenseada com o Questionário da Amostra, no estrato de setores urbanos de aglomerados subnormais..

b) Para variáveis contínuas:

$y_{hij} = valor\ da\ variável\ de\ estudo\ na\ unidade\ j\ do\ setor\ i\ no\ estrato\ h.$

Cálculo das estimativas de precisão das estimativas pontuais da Tabulação Avançada

Para avaliar a precisão das estimativas de totais, foram calculados os respectivos Coeficientes de Variação (CVs), definidos como:

$$cv(\hat{Y}) = \frac{\sqrt{v(\hat{Y})}}{\hat{Y}};$$

onde $v(\hat{Y})$ é a estimativa da variância da estimativa de total $\hat{Y}$, que foi calculada conforme fórmulas definidas no capítulo 8 de Särndal, Swensson e Wretman (1992).

A partir do coeficiente de variação, pode-se construir um intervalo de confiança, para um total em questão, dado pela expressão:

$$\hat{Y} - z_\alpha \times \hat{Y} \times cv(\hat{Y}) \leq Y \leq \hat{Y} + z_\alpha \times \hat{Y} \times cv(\hat{Y});$$

onde:

Y é o valor verdadeiro do total em questão;

$\hat{Y}$ é a estimativa amostral do total;

$cv(\hat{Y})$ é o coeficiente de variação da estimativa;

z_α é o valor da ordenada da distribuição normal padrão para um nível á de significância.

Como nas tabelas divulgadas na Tabulação Avançada existe um número muito grande de estimativas de total, optou-se por não colocar os valores estimados dos CVs, e sim uma letra ao lado de cada estimativa pontual, correspondente a uma faixa de valores para o coeficiente de variação. As faixas utilizadas foram as sugeridas em Albieri (1999a), apresentadas no Quadro 12.1, a seguir.

Quadro 12.1 - Indicadores de faixas de coeficiente de variação utilizados nas tabelas da Tabulação Avançada do Censo Demográfico 2000

Indicador	Faixas de CV (%)
z	zero
a	de 0,0 até 0,5
b	mais de 0,5 até 1,0
c	mais de 1,0 até 2,5
d	mais de 2,5 até 5,0
e	mais de 5,0 até 7,5
f	mais de 7,5 até 10,0
g	mais de 10,0 até 15,0
h	mais de 15,0 até 25,0
i	mais de 25,0 até 35,0
j	mais de 35,0 até 50,0
k	mais de 50,0

A partir das letras indicativas dos valores dos coeficientes de variação, é possível calcular um intervalo de confiança aproximado para o total desejado, usando os limites das faixas de valores. Exemplificando, pode-se construir um intervalo de confiança de 95% para um dado total, cujo coeficiente de variação da estimativa esteja na faixa b usando o limite superior dessa faixa, ou seja:

$$\hat{Y} - 1{,}96 \times \hat{Y} \times 0{,}01 \leq Y \leq \hat{Y} + 1{,}96 \times \hat{Y} \times 0{,}01$$

Tratamento dos dados da Tabulação Avançada

Convém ressaltar que uma vez selecionados os setores que pertencerem à amostra da tabulação, foram definidos procedimentos para que a sua apuração fosse realizada de forma prioritária, ou seja, esses setores tiveram prioridade nas primeiras etapas de apuração dos questionários, desde a remessa para os centros de captura, passando por todos os procedimentos relacionados com a própria captura, a saber, leitura ótica, verificação e crítica de quantidades. Todos os demais procedimentos de apuração dos questionários pertencentes aos setores da amostra da tabulação avançada foram realizados de forma separada e independente da apuração para a obtenção dos resultados definitivos. Para tanto, a base de dados desses questionários foi duplicada e neles foram aplicados apenas os procedimentos de validação para a geração dos resultados preliminares divulgados. Todas as etapas de apuração, após a captura, não foram incorporadas ao processo definitivo.

Excetuando as informações referentes à estrutura familiar e gênero, os dados divulgados para os demais temas da publicação não passaram pelo processo de crítica eletrônica, adotado na divulgação de resultados definitivos, e que tem por finalidade eliminar eventuais inconsistências entre as informações dos diversos quesitos do questionário, que podem ter origem na coleta de dados ou na fase de reconhecimento de marcas e caracteres.

Portanto, as informações que apresentaram alguma inconsistência, do tipo quesito omitido quando deveria estar preenchido, erro de seqüência no preenchimento do questionário ou impossibilidade de alocação de um valor em alguma célula de tabela, foram incluídas apenas nas colunas ou linhas de total. Assim, os valores dos totais de linhas e colunas não necessariamente coincidem com as somas dos valores das parcelas correspondentes.

Conteúdo da publicação da Tabulação Avançada

A publicação *Tabulação avançada do censo demográfico 2000: resultados preliminares da amostra* (2002) contém, para o total do País e para as grandes regiões e unidades da federação, um total de 38 tabelas com características dos domicílios e das pessoas, captadas por meio do questionário da amostra sobre os seguintes temas: características gerais da população, educação, migração, nupcialidade, trabalho, famílias e domicílios.

A publicação incluiu, além das notas metodológicas, comentários dos resultados sobre as características gerais: cor ou raça, religião e deficiência; características da educação, migração, nupcialidade, famílias e domicílios. Incluiu também um CD-ROM com as 38 tabelas divulgadas.

12.4.2 Fecundidade e mortalidade Infantil

No tocante aos temas Fecundidade e Mortalidade Infantil, a divulgação dos resultados preliminares da amostra teve um tratamento diferenciado dos demais temas que compuseram a *Tabulação avançada do censo demográfico 2000: resultados preliminares da amostra*. Em vez de se utilizar uma subamostra de 0,24% dos domicílios, o IBGE optou por uma apresentação em separado da Tabulação Avançada utilizando-se, para fins de geração das correspondentes estimativas, toda a amostra do censo. Esta foi uma decisão baseada nas especificidades das variáveis envolvidas que guardam certo grau de complexidade na coleta das informações e, sobretudo, porque os indicadores derivados são, pela sua natureza, bastante sensíveis a flutuações amostrais.

Além disso, a elaboração das estimativas dos parâmetros da Fecundidade e da Mortalidade Infantil requer a aplicação de metodologias específicas. Para maiores esclarecimentos, podem ser consultados: Brass (1971), Brass et al. (1968, 1975), Camisa (1975), Coale e Trussell (1974), Oliveira (1991) e Trussell (1975). Cabe ressaltar, também, que nesta etapa de divulgação dos resultados preliminares da amostra do Censo Demográfico 2000 fez-se uso das informações obtidas após o processo de captura de dados, portanto, ainda não submetidas à crítica quanto à sua consistência.

12.4.2.1 Cálculo das estimativas

Para o cálculo das estimativas da Fecundidade e da Mortalidade Infantil foram introduzidos alguns filtros que tinham como objetivo fornecer uma visão, ainda que muito preliminar, dos dados com algum tipo de crítica. Isso foi levado a efeito por tema, considerando-se os registros com erro de preenchimento e de não resposta referentes às mulheres, para as seguintes variáveis:

- Fecundidade: filhos tidos nascidos vivos, filhos sobreviventes ou filhos tidos nos últimos doze meses anteriores ao censo demográfico 2000, e

- Mortalidade Infantil: filhos tidos nascidos vivos ou filhos sobreviventes.

Uma descrição detalhada da implementação desses filtros pode ser encontrada no anexo do CD-ROM que acompanha a metodologia.

São descritos a seguir os pressupostos relativos às técnicas utilizadas para o cálculo das taxas de fecundidade e de mortalidade infantil.

a) Fecundidade

No tocante à Fecundidade, a técnica empregada é a metodologia proposta por Brass (1971), que consiste em combinar três informações:

- mulheres em idade fértil (de 15 a 49 anos de idade), classificadas por grupos qüinqüenais de idade [M (i); i = 1,2,3,....,7],

onde $i = 1 \rightarrow$ 15 a 19 anos

$i = 2 \rightarrow$ 20 a 24 anos

$i = 3 \rightarrow$ 25 a 29 anos

$i = 4 \rightarrow$ 30 a 34 anos

$i = 5 \rightarrow$ 35 a 39 anos

$i = 6 \rightarrow$ 40 a 44 anos

$i = 7 \rightarrow$ 45 a 49 anos

- filhos tidos nascidos vivos, declarados pelas mulheres, classificados segundo os mesmos grupos qüinqüenais de idade das mulheres [FNV (i)], e

- filhos tidos nascidos vivos nos 12 meses anteriores à data de referência do censo, declarados pelas mulheres, tabulados segundo os grupos qüinqüenais de idade das mulheres de 15 a 49 anos [FUA (i)], sendo esta uma informação derivada do quesito "data de nascimento do último filho tido nascido vivo".

Tais informações foram obtidas das mulheres de 10 anos ou mais de idade, mas a técnica foi concebida para ser aplicada ao contingente feminino de 15 a 49 anos.

Inicialmente, são calculadas:

- as parturições médias, P(i), que: representam a fecundidade retrospectiva acumulada das mulheres, obtidas mediante a seguinte relação:

P (i) = FNV (i) / M (i), i = 1, 2, ... , 7, e

- as taxas específicas de fecundidade por grupos de idade das mulheres, que expressam a fecundidade atual das mulheres, através do seguinte cálculo:

f (i) = FUA (i) / M (i), i = 1, 2, ... ,7.

Em síntese, a aplicação da técnica consiste na comparação das parturições [P(i)] com a fecundidade atual acumulada [F(i)], obtida dos nascimentos dos últimos 12 meses. Esta comparação é feita mediante a análise da série P (i) / F (i), da qual será extraído o fator de correção das taxas de fecundidade atuais f (i).

Isso porque, para a aplicação da técnica, também conhecida como técnica da razão P/F, o requerimento básico é a aceitação de duas hipóteses relacionadas com a informação de referência. A primeira diz que a estrutura por grupos de idade das taxas de fecundidade atual é aceitável, ainda que não seja o nível estimado a partir dela, isto é, a Taxa de Fecundidade Total. A segunda hipótese faz referência à melhor qualidade da informação sobre a fecundidade retrospectiva, associada aos grupos 20 a 24 e 25 a 29 anos de idade. A depender do caso específico, esses grupos etários são empregados como bons indicadores do nível da fecundidade.

A técnica de Brass traz implícitos, porém, dois pressupostos metodológicos que relacionam f(i) e P (i):

- que as mulheres sobreviventes à data do censo são regidas pela mesma lei de fecundidade das mulheres que já faleceram, e

- que a fecundidade tenha permanecido constante ao longo do tempo. As mulheres chegam a uma parturição média correspondente ao grupo 45 a 49 anos de idade expostas aos mesmos riscos de fecundidade prevalecentes nos últimos 12 meses.

Se estes pressupostos se cumprem, concomitantemente com a inexistência de erros de declaração da idade e do número de filhos, a razão P (i) / F (i) = 1.

Mas, em populações reais, sobretudo quando a fecundidade experimenta declínios, tal relação costuma ter um comportamento crescente à medida que aumenta a idade das mulheres, e assume valores superiores à unidade. O fator

de correção das taxas de fecundidade atuais será selecionado a partir da experiência reprodutiva das mulheres mais jovens, das de 20 a 24 anos ou das de 25 a 29 anos de idade, respectivamente, P (2) / F (2) ou P (3) / F (3).

b) Mortalidade infantil

Já no que concerne à Mortalidade Infantil, a técnica empregada foi a variante de Trussell (1975), da técnica originalmente proposta por Brass (1971).

A informação básica necessária para estimar a mortalidade infantil é a seguinte:

- mulheres em idade fértil (de 15 a 49 anos de idade), classificadas por grupos qüinqüenais de idade [M (i); i = 1 = 15 a 19 anos, i = 2 = 20 a 24 anos,....., i = 7 = 45 a 49 anos],

- filhos tidos nascidos vivos, declarados pelas mulheres, classificados segundo os mesmos grupos qüinqüenais de idade das mulheres [FNV (i)], e

- filhos sobreviventes, declarados pelas mulheres, classificados segundo os mesmos grupos qüinqüenais de idade das mulheres [FV (i)]. Esta informação está referida ao momento do censo.

Com estes dados podem ser calculadas, inicialmente, as proporções de filhos falecidos com respeito ao total de filhos nascidos vivos, segundo a idade das mulheres:

$$D(i) = 1 - [\, FV(i) / FNV(i) \,]$$

onde D (i), por si só, constitui uma medida da mortalidade, mas tem a limitação de não ser um indicador convencional, por estar referido à idade das mulheres e não à idade dos filhos. Nesse sentido, Brass desenvolveu um procedimento que permite transformar as proporções D (i) em medidas convencionais de mortalidade nos primeiros anos de vida. O autor demonstrou haver uma relação empírica entre D (i) e a probabilidade de morte desde o nascimento até uma idade exata x, Q (x). A relação entre estas medidas se estabelece mediante as seguintes relações:

Quadro 12.2 – Correspondência entre D(i) e Q(x)

i	Q (x)	=	K (i) * D (i)
1	Q (1)	=	K (1) * D (1)
2	Q (2)	=	K (2) * D (2)
3	Q (3)	=	K (3) * D (3)
4	Q (5)	=	K (4) * D (4)
5	Q (10)	=	K (5) * D (5)
6	Q (15)	=	K (6) * D (6)
7	Q (20)	=	K (7) * D (7)

K (i) é um fator muito próximo a 1 (um), o que permite transformar as D (i) em Q (x). Brass calculou um conjunto de valores de K (i) com base em um modelo teórico no qual intervêm uma função de fecundidade e uma lei de mortalidade. Foi demonstrado que, neste modelo, os multiplicadores dependem

principalmente da estrutura por idade da fecundidade, no sentido de que, quanto mais cedo as mulheres tiverem seus filhos, maior será o tempo de exposição ao risco de morte de seus filhos. Por esse motivo, os parâmetros de entrada para a obtenção dos valores de K(i) são indicadores dessa estrutura, P (1) / P (2) e P (2) / P (3), sendo P (i) a parturição média das mulheres no grupo etário i (i = 1 para 15 a 19, i = 2 para 20 a 24, etc.).

Cada Q (x) estimada corresponde a momentos distintos anteriores à data de referência do censo. Na medida em que se avança na idade das mulheres, a estimativa corresponde a um passado mais distante. Feeney (1976, 1980) foi o primeiro a desenvolver idéias a respeito de como localizar as estimativas no tempo. Descobriu que, ao supor um declínio linear, qualquer que seja a intensidade do declínio, a mortalidade é a mesma num momento de tempo anterior ao censo. A partir dessa idéia, Coale e Trussell (1977) desenvolveram um procedimento para determinar os valores de t (x) (número de anos anteriores ao censo) para cada Q (x) estimada. Baseando-se nos modelos de fecundidade de Coale e Trussell e nas quatro famílias (Norte, Sul, Leste e Oeste) de tábuas-modelo de mortalidade de Coale e Demeny (1966), Trussell elaborou quatro conjuntos de regressões para o cálculo de K (i) e t (x).

Com o propósito de se obter uma medida comparável no tempo, usando-se as tábuas-modelo de Coale e Demeny (1966), as tábuas-modelo Brasil de Frias e Rodrigues (1981) ou uma transformação logital, todas as Q (x) obtidas, mediante o emprego da técnica de Trussell, foram transformadas em Q (1), ou seja, em probabilidades de um recém-nascido falecer antes de completar o primeiro ano de vida, devidamente localizadas no tempo.

A técnica de Trussell foi aplicada às informações provenientes dos Censos Demográficos 1970, 1980, 1991 e 2000, formando um conjunto de estimativas de Q (1), que compreendia o período de 1960 - 1998. Tais estimativas foram suavizadas, mediante médias móveis, de maneira a eliminar possíveis flutuações que normalmente existem, derivadas, principalmente, da má declaração, por parte das mulheres, quanto ao número de filhos nascidos vivos e sobreviventes. A esta série, já suavizada, ajustou-se uma função logística. Deve-se esclarecer que se teve o cuidado para que os ajustes realizados não implicassem diferenças significativas dos valores observados, especialmente nos anos próximos a 2000, o que foi possível, simulando-se valores para as assíntotas inferior e superior da função logística. Ao proceder assim, pouca variabilidade foi encontrada entre os valores observados e ajustados, ao longo do período considerado. Isto proporcionou um ajuste bastante satisfatório das probabilidades de morte no primeiro ano de vida, possibilitando realizar projeções das mesmas para o ano 2000.

A fim de realizar as análises com conhecimento do significado das estimativas, são os seguintes os pressupostos implícitos da técnica utilizada:

- que a fecundidade tenha permanecido constante num passado recente. Segundo Feeney (apud MANUAL X...,1983), as estimativas são suficientemente robustas de modo que os desvios não têm importância se não se cumpre esse suposto,

- que a mortalidade na infância tenha uma evolução linear através do tempo,

- que as leis de mortalidade e fecundidade usadas no modelo representem as mesmas condições da população em estudo,

- que não haja associação entre a mortalidade das mulheres (mães) e de seus filhos. Obviamente, não se tem informação sobre a mortalidade dos filhos cujas mães já faleceram, e, no caso em que sua mortalidade fosse maior que a dos filhos com mães vivas, as estimativas da mortalidade nos primeiros anos de vida estaria subestimada, e

- que não exista associação entre a mortalidade infanto-juvenil e a idade das mulheres (mães).

Além destes pressupostos, a informação básica deve cumprir certas condições:

- que não haja omissão diferencial na declaração do número de filhos nascidos vivos e sobreviventes,

- que não haja mortalidade diferencial entre os filhos das mulheres que declaram e as que não declaram a informação, e

- que a declaração da idade das mulheres seja correta.

As taxas estimadas segundo esses critérios são apresentadas a seguir.

Tabela 12.16 - Taxas de Fecundidade Total e Taxas de Mortalidade Infantil com base em dados censitários - Brasil - 1980-2000

Ano	Taxa de fecundidade total	Taxa de mortalidade infantil (por 1 000 nascidos vivos)
1980	4,4	82,6
1991	2,9	47,7
2000	2,3	28,3

Fonte: IBGE, Censo Demográfico 1980/2000.

c) Conceitos e definições

No Censo 2000, os quesitos do bloco de fecundidade foram indagados a todas as mulheres com 10 anos ou mais de idade, na data de referência do censo, ou seja, nascidas até 31 de julho de 1990.

- Filhos tidos nascidos vivos até 31 de julho de 2000 – Considerou-se como filho tido nascido vivo aquele que, após a expulsão ou extração completa do corpo da mãe, independentemente do tempo de duração da gravidez, manifestou qualquer sinal de vida (respiração, choro, movimentos de músculos de contração voluntária, batimento cardíaco, etc.), ainda que tenha falecido em seguida. O número de filhos tidos nascidos vivos foi registrado segundo o sexo.

- Filhos tidos que estavam vivos em 31 de julho de 2000 – O número de filhos tidos que estavam vivos em 31 de julho de 2000 foi registrado segundo o sexo.

- Sexo do último filho tido nascido vivo até 31 de julho de 2000. As opções de resposta foram: masculino ou feminino.

- Data de nascimento ou idade presumida do último filho tido nascido vivo até 31 de julho de 2000 – Registrou-se o mês e o ano de nascimento do último filho tido nascido vivo até 31 de julho de 2000. Se, esgotados todos os esforços, não fosse possível a obtenção do mês e ano de nascimento do último filho tido nascido vivo, registrou-se sua idade presumida, fornecida pela pessoa entrevistada.

- Sobrevivência do último filho tido nascido vivo até 31 de julho de 2000 – As opções de resposta foram: sim; não; não sabe.

- Filhos tidos nascidos mortos – São os óbitos ocorridos de todo o produto da concepção, a partir da 28a semana de gestação, antes de sua extração ou expulsão completa do corpo da mãe. A informação foi coletada segundo o sexo.

- Taxa de Fecundidade Total – Expressa o número de filhos que, em média, teria uma mulher, pertencente a uma coorte hipotética de mulheres, que durante sua vida fértil tiveram seus filhos de acordo com as Taxas de Fecundidade, por Idade do período em estudo e que não estiveram expostas a riscos de mortalidade desde o nascimento até o término do período fértil.

- Taxa Específica de Fecundidade por idade – É geralmente calculada por grupo qüinqüenal de idade, desde os 15 até os 49 anos. A taxa resulta da divisão do número de filhos nascidos vivos de mulheres do grupo de idade, em um período de tempo próximo à data do Censo Demográfico, usualmente os últimos 12 meses, pelo total de mulheres do mesmo grupo etário.

- Coorte – Conjunto de indivíduos que estão experimentando um acontecimento similar no transcurso de um mesmo período de tempo.

- Coorte hipotética de mulheres – Num censo demográfico, a classificação das mulheres por grupos qüinqüenais de idade, dentro do período fértil, está associada a uma análise de período. Uma análise de coorte considera, por exemplo, um grupo de mulheres que ingressa no período fértil e, ao longo do tempo, observa-se o comportamento do mesmo frente aos riscos de procriação. Entretanto, em um único censo demográfico, mesclam-se distintas gerações de mulheres e, de acordo com o conceito da Taxa de Fecundidade Total, supõe-se o acompanhamento de como essas mulheres vão tendo seus filhos ao longo do tempo. Por esse motivo, na definição conceitual da Taxa de Fecundidade Total, é necessário enfatizar que o grupo de mulheres em questão trata-se de uma coorte hipotética.

- Taxa de Mortalidade Infantil – É definida como o número de óbitos de menores de 1 ano de idade (por mil nascidos vivos), em determinada área geográfica e período, e interpreta-se como a estimativa do risco de um nascido vivo morrer durante o seu primeiro ano de vida.

d) Expansão da amostra e divulgação dos resultados

É importante assinalar que os pesos preliminares para a expansão da amostra, que viabilizaram os cálculos dos indicadores de Fecundidade e Mortalidade Infantil, foram obtidos através do inverso da fração de amostragem, observada no setor censitário, e calibrados de tal forma que as estimativas de total de pessoas por sexo se igualassem aos valores correspondentes do Conjunto Universo – que compreende o conjunto de características básicas investigadas para o total da população e dos domicílios, em cada Unidade da Federação. Os pesos, assim determinados, foram atribuídos a cada domicílio; todas as pessoas residentes em um mesmo domicílio receberam peso idêntico ao do domicílio.

Os níveis geográficos para a divulgação dos Resultados Preliminares dos temas Fecundidade e Mortalidade Infantil foram o Brasil como um todo e suas cinco grandes regiões. Para o tema Fecundidade foram divulgados indicadores representativos do nível – de 1940 até 2000 – e do padrão etário – de 1980 até 2000. O nível da Fecundidade está representado pela Taxa de Fecundidade Total e o padrão etário pelas Taxas Específicas de Fecundidade por Idade. No caso da Mortalidade Infantil, foram apresentadas as respectivas séries históricas das Taxas de Mortalidade Infantil, abrangendo o período de 1990 a 2000.

12.5 Resultados da amostra

Comparativamente ao processo de apuração dos Resultados do Universo, a tarefa referente aos dados da amostra, pelo maior volume do trabalho de crítica e imputação e por incorporar as tarefas da Codificação e Expansão, apresenta-se bem mais intensa. A seguir, são tratadas cada uma dessas atividades.

12.5.1 Codificação

Tendo como referência o Censo de 1991, a Codificação pode ser apontada como uma das partes da apuração dos dados que incorporou um grande número de modificações para o Censo. A maior novidade foi a implantação de uma rotina de aplicação de códigos – codificação propriamente dita – reformulada, que tinha como objetivo garantir a qualidade do trabalho, com um prazo de execução bastante reduzido.

Assim, nesse contexto de procura de maior eficiência, foi também estabelecida a etapa de Verificação, que compreendeu um conjunto de ações que buscavam aprimorar a atividade de aplicação de códigos, vindo a constituir-se numa inovação no processo de trabalho da apuração. Entenda-se, portanto, a fase de Codificação do Censo Demográfico 2000, compreendendo duas fases: Aplicação de Códigos e Verificação de Códigos.

12.5.1.1 Formação dos lotes

Para a formação dos lotes a serem trabalhados, na Codificação e nas atividades de Crítica Intra-Registros das informações do CD 1.02, foram considerados, em boa parte, os critérios utilizados para os dados do Conjunto Universo. Assim, primeiramente, os setores foram classificados/ordenados levando-se em conta a Unidade da Federação, a situação do domicílio, a mesorregião, a microrregião, o município, o distrito e o subdistrito.

A diferença em relação ao Conjunto Universo ficou por conta do processo de escolha dos lotes. Partindo do mesmo limite superior de 90000 domicílios particulares ocupados – DPO –, foram definidos os lotes, buscando-se respeitar os vários níveis da classificação.

Para as Unidades da Federação, onde o total de DPO não ultrapassou os 90 000 domicílios, para cada situação urbana e rural, os lotes foram formados com todos os setores dessas UFs; caso contrário, cada lote foi composto pela mesorregião. As áreas urbanas dos municípios de Salvador, Belo Horizonte e Porto Alegre formaram, cada uma, um lote exclusivo. Para a área urbana do município de São Paulo, foram formados lotes através de áreas pré-definidas, agrupando-se os seus distritos. No caso do município do Rio de Janeiro, foram formados lotes através de grupamentos de subdistritos.

A tabela seguinte apresenta algumas informações para os 215 lotes que foram formados:

Tabela 12.17 - Número de lotes e de domicílios particulares ocupados e de pessoas recenseadas, informados no SIGC, referentes aos Questionários da Amostra, segundo as Unidades da Federação

Unidades da Federação	Número de lotes	DPO	Pessoas
Brasil	**215**	**5 247 272**	**20 199 963**
Rondônia	2	42 966	171 504
Acre	2	16 711	70 878
Amazonas	2	63 694	314 039
Roraima	2	9 663	41 621
Pará	3	145 106	689 625
Amapá	2	11 653	55 192
Tocantins	2	42 648	175 455
Maranhão	2	149 698	701 509
Piauí	2	94 195	404 761
Ceará	14	200 174	864 021
Rio Grande do Norte	2	92 118	388 864
Paraíba	2	116 895	486 823
Pernambuco	10	223 992	932 492
Alagoas	2	77 516	347 126
Sergipe	2	54 825	230 323
Bahia	15	376 477	1 593 721
Minas Gerais	25	608 183	2 339 334
Espírito Santo	2	98 067	368 075
Rio de Janeiro	19	437 509	1 503 726
São Paulo	40	1 118 165	4 017 990
Paraná	20	332 772	1 215 472
Santa Catarina	12	191 741	69 139
Rio Grande do Sul	15	361 390	1 206 575
Mato Grosso do Sul	2	68 316	25 017
Mato Grosso	2	85 402	325 034
Goiás	10	173 456	614 801
Distrito Federal	2	53 940	199 442

Fonte: IBGE, Censo Demográfico 2000, Sistema de Indicadores Gerenciais da Coleta.

12.5.1.2 Modelo de codificação automática e assistida

O sistema de codificação automática/assistida, disponível no IBGE para ser utilizado por pesquisas e censos, adotava um modelo idealizado em 1988 por ocasião da realização do Censo Experimental daquele ano. Esse modelo serviu de base para a confecção do sistema utilizado na codificação do Censo Demográfico 1991 e pelas PNADs da última década. Ao longo desses anos, o sistema foi aprimorado em suas funções, porém, nenhuma revisão foi efetuada no seu modelo conceitual.

Modelo utilizado no Censo Demográfico 1991

O modelo utilizado no Censo Demográfico 1991, que tinha sido implementado num sistema de codificação automática/assistida, apresentava, resumidamente, as características descritas a seguir:

a) para cada quesito a codificar, era necessário montar um arquivo (arquivo de descritores) com as várias descrições possíveis para o quesito em questão, contendo o código correspondente. A partir desses arquivos, era criado um banco de códigos contendo, para cada quesito a ser codificado, as descrições correspondentes e as palavras em sua forma normal e fonética e os códigos associados. Esta etapa preparatória era realizada, de forma centralizada, somente uma vez antes de iniciar a codificação e poderia sofrer atualizações, de forma centralizada, dependendo das necessidades surgidas ao longo do processo de codificação;

b) o processo de codificação era realizado sobre um arquivo magnético com os dados dos questionários capturados (via digitação ou reconhecimento ótico de caracteres);

c) a comparação do texto oriundo do quesito com os textos das descrições era feita palavra a palavra. Ao se dividir os textos em palavras, eram eliminadas as preposições e os artigos, além das palavras de uma lista opcional a serem eliminadas, fornecida pelo responsável pela codificação do quesito. O objetivo desta lista era possibilitar a eliminação de palavras que não contribuiriam para a codificação do texto (ex.: a palavra igreja na codificação do quesito religião). Na divisão do texto em palavras, os sinônimos eram, também, considerados. Toda palavra, num conjunto de palavras consideradas sinônimos, era convertida para uma palavra padrão. As palavras também sofriam uma transformação fonética que consistia, em resumo, na eliminação do plural, do gênero (masculino e feminino) e na substituição de uma letra por outra do mesmo som (esta técnica auxilia no reconhecimento das palavras e resolve também alguns erros de grafia ou de digitação). O método utilizado para tratamento do texto a ser codificado foi o mesmo utilizado para geração do banco de códigos a partir dos arquivos de descritores.

d) a aplicação do código era feita a partir da comparação dos textos obtidos dos questionários com os textos armazenados no banco de códigos, com objetivo de atribuir ao texto um código numérico. Para esta comparação, o texto do questionário era dividido em palavras e cada palavra pesquisada em sua forma normal. Caso não fosse encontrada, era feita a pesquisa em sua forma fonética. No caso do texto conter mais de uma palavra, ou uma palavra não reconhecida, esta era pesquisada no contexto das descrições, onde apareciam as outras palavras, podendo ser feitas sugestões de palavras semelhantes. Reconhecidas as palavras, eram pesquisadas nas descrições onde apareciam, podendo surgir três situações:

d.1) uma única descrição era encontrada: codificação automática;

d.2) mais de uma descrição era encontrada: o codificador devia escolher uma das descrições apresentadas;

d.3) nenhuma descrição era encontrada: o codificador digitava um novo texto.

O modelo de codificação implementado apresentava claramente duas etapas de codificação: uma primeira chamada de codificação automática ou pré-codificação (*batch*) e uma segunda chamada de codificação assistida ou complementar (*on-line*). Na codificação automática eram codificados, questionário a questionário, aqueles textos para os quais era encontrado um único código no banco de códigos (d.1), o que determinava o fim do processo de codificação para o quesito do questionário. Na codificação assistida, um codificador fazia, *on-line*, a escolha do código, entre os apresentados com base na descrição (d.2) ou digitava um novo texto para o quesito em questão (d.3).

Para todos os quesitos a codificar de um questionário, o sistema realizava a codificação automática na seqüência em que os quesitos estavam no questionário. O codificador atuava somente após a execução da pré-codificação (processamento *batch*) para completar a tarefa nos questionários com codificação pendente, obedecendo a ordem dos quesitos no questionário. É importante registrar que um mesmo codificador era responsável pela codificação assistida dos diferentes quesitos (ocupação, atividade, religião etc.).

Avaliação do modelo

Uma das maiores atividades no processamento de um censo é a codificação, talvez só inferior à captura de dados.

O modelo até então disponível era, naturalmente, superior à codificação manual, processo lento que requer um grande contingente de pessoas com perfil adequado e consome muito tempo.

O esquema utilizado no modelo disponível (separar o texto em palavras, eliminar algumas palavras, utilizar palavra padrão – sinônimo, abreviatura – transformação fonética, e criar o banco de códigos, levando em conta a freqüência das palavras e a forma de busca do código a ser associado), com certeza, trouxe maior eficiência ao processo.

Se o conjunto de descrições encontradas no Censo Demográfico 1991 para os distintos quesitos a codificar fosse incorporado aos arquivos de descritores, com certeza, aumentaria bastante a eficiência e a qualidade da codificação do Censo Demográfico 2000, utilizando-se o modelo disponível.

No modelo descrito, pode-se verificar que muito investimento foi feito na "máquina para obter um código a partir de um texto". O modelo chegou a um ponto em que, apesar de ser possível melhorar o custo/benefício, as modificações seriam difíceis de implementar e os ganhos pouco perceptíveis.

O que chamou a atenção dos responsáveis pela codificação do Censo Demográfico 2000 não foi a eficiência dessa máquina, mas a forma como foi utilizada.

Nesse modelo, a codificação era feita questionário a questionário. Tanto na codificação automática, como na assistida, os quesitos eram trabalhados, um a um, na seqüência em que apareciam no questionário. O codificador tratava todos os quesitos do questionário, seqüencialmente. Isto fez pensar o seguinte: criou-se uma bela "máquina" de codificar, mas que estava sendo utiliza-

da, simulando um processo de codificação manual. Nesse processo, um codificador ia retirando questionários de uma pilha e codificando os quesitos, um a um, sem precisar recorrer ao velho e tradicional manual de códigos (transformado no banco de códigos). Automatizou-se um processo, porém, não se o informatizou. Isto é, as tarefas manuais foram transferidas para o computador, retirando-se uma parte do trabalho manual do codificador. A outra parte desse trabalho continuava sendo feita manualmente, porém, assistida pelo computador, sem que nenhuma informação gerada durante o processo fosse utilizada para aumentar a eficiência do mesmo. Desta forma, a tarefa mais difícil era resolvida por uma única pessoa, que devia codificar todos os quesitos do questionário, uma vez que a "inteligência" colocada na "máquina" não conseguia resolver a codificação de forma automática. Por outro lado, a partir da existência de um sistema de codificação automática/assistida, criou-se a idéia de que com um pequeno treinamento (basicamente no uso do sistema) uma pessoa estaria, rapidamente, habilitada a codificar qualquer quesito.

A codificação apresenta diferentes graus de dificuldade que variam de um quesito para outro. Os índices de codificação automática do Censo Demográfico 1991, para quesitos religião (da ordem de 80%), Unidade da Federação ou país estrangeiro de nascimento (80%), município (80%), curso concluído (79%) são totalmente diferentes ,se comparados com alcançados em ocupação (cerca de 20%) e atividade (inferior a 40%).

Religião, local de nascimento, município e curso concluído causam menos problemas, em virtude de serem mais facilmente entendidos e apresentarem um conjunto de possibilidades bem limitado. Os maiores problemas resumem-se a erros de grafia, uso de sinônimos, abreviaturas das mais diversas formas, etc.

Por outro lado, os quesitos ocupação e atividade apresentam uma grande variedade de descrições, às vezes feitas de uma forma muito geral para que o sistema as codifique. Estes quesitos podem depender de conceitos e entendimentos muitas vezes restritos à esfera de conhecimento de especialistas no assunto, o que não se pode esperar de codificadores que passam por um pequeno período de treinamento. Este é, com certeza, o ponto mais crítico desse modelo.

Modelo utilizado no Censo Demográfico 2000

O modelo utilizado para aplicação de códigos no Censo Demográfico 2000 foi uma combinação entre a aplicação manual e a automática. Representa algo como uma pós-codificação, em que os códigos para cada quesito só foram atribuídos após se ter conhecimento de todas as descrições encontradas nos questionários de um determinado conjunto (lote de codificação).

A aplicação de códigos foi feita em cada um dos lotes constituídos para as descrições registradas no Questionário da Amostra – CD 1.02 – correspondentes às informações coletadas em aberto, para os quesitos Religião, Migração, Curso, Ocupação e Atividade. O tema Migração era constituído por 5 quesitos de codificação, a saber: “Qual é a unidade da federação ou país estrangeiro de nascimento?” “Qual é a unidade da federação ou país estrangeiro de residência anterior?” “Em que município residia em 31 de julho de 1995?” “Em que unidade da federação ou país estrangeiro residia em 31 de julho de 1995?”, e “Em que município e unidade da federação, ou país estrangeiro, trabalha ou estuda?”. Compreende-se assim que, exceto em Migração, tema e quesito de codificação sejam a mesma coisa, e que foram codificados 9 quesitos em cada um dos 215

lotes trabalhados. O tamanho do lote de codificação foi determinado em função do lote preparado para as etapas de codificação e crítica.

O modelo apresentou, para cada quesito a ser codificado, as seguintes características:

- as descrições encontradas nos questionários foram armazenadas para cada pessoa e transferidas para uma base de dados em separado. Nesta base, as informações semelhantes, para cada tema, foram unificadas e grupadas, segundo os critérios automáticos de aplicação de códigos. As descrições foram grupadas por semelhança (através de palavras-chave que determinavam o código ou códigos associados usando a "máquina");

- a aplicação de códigos foi feita, nesta base de descrições, por lote. Todos os registros desta base, com as mesmas características, foram "codificados" de uma única vez, apresentando a freqüência de descrições que eles representavam. O sistema atuou sempre com controle de freqüência, de ocorrências das descrições, individuais e grupadas no lote;

- a aplicação de um código era, inicialmente, confirmada pelo codificador do tema, mediante a validação do código proposto pelo sistema. Depois de atingir uma determinada freqüência acumulada e após a validação do orientador, a aplicação do código passava a ser feita de forma automática, sem precisar da confirmação do codificador; e

- terminada a aplicação de códigos do arquivo de descrições básicas do lote, os arquivos originais (lote de codificação), contendo registros individualizados e o arquivo com as descrições já codificadas, passaram por um batimento para atribuição do código aos registros individuais no lote.

As vantagens do modelo utilizado em relação ao modelo anterior, entre outras, são: maior automação do processo de aplicação de códigos, uniformidade do processo e utilização de codificadores especializados em cada tema.

Etapas do Sistema

O sistema implementado era composto pelas seguintes etapas:

Extração de textos

O sistema lia o arquivo com o lote a codificar e extraía dos registros de pessoas, os textos referentes ao quesito (tema) selecionado, guardando-os com sintaxe única, em arquivos do banco com a freqüência com que ocorreiam. Em seguida, os textos eram verificados no banco de códigos, separando os textos corretos (aqueles que tem codificação única ou múltipla) e os não corretos (aquele para os quais não existe código) em relação ao descritor. Criavam-se,assim, três grupos de textos: os que codificavam, os que tinham codificação múltipla, e aqueles para os que não haviam códigos associados. Durante a extração, o sistema contabilizava totais de domicílios e pessoas lidas, quantidade de textos extraídos e quantos textos distintos existiam (textos diferentes). Ao final do processo, eram apresentadas as estatísticas da extração.

Correção de caracteres

O sistema lia o arquivo com os textos extraídos. Os textos que não tinham código associado (codificação não é única nem múltipla) eram exibidos em caixa de lista, permitindo sua correção; se um texto em trabalho já tivesse passado por correção anteriormente, o sistema sugeria a sua última correção feita. Se a freqüência de alguma correção atingisse um valor determinado como limite e houvesse autorização do orientdor, era feita automaticamente (aprendizado através do uso). A freqüência para correção automática era única para o quesito durante o processamento dos lotes.

Durante esta etapa, os textos, corrigidos ou não, eram mantidos em um arquivo de correções para serem pesquisados no arquivo descritor. Essa pesquisa era feita sempre ao iniciar a correção. Caso a correção de caracteres fosse re-executada, o sistema fazia a pesquisa no arquivo descritor, retornando como errados aqueles que não tinham sido encontrados. Caso a correção de caracteres fosse concluída sem re-execução, este batimento era feito na próxima etapa (Agrupamento por códigos) e os textos não encontrados eram agrupados como ainda não codificados.

Agrupamento por códigos

Os textos corrigidos eram codificados dentro dos grupos em cada tema em códigos únicos, múltiplos e não codificados, e disponibilizados para confirmação. Durante o agrupamento, era feita uma contagem para cada uma dessas situações, sendo as estatísticas apresentadas ao final do processo.

Confirmação de códigos

Os textos grupados eram exibidos em uma lista para confirmação. Se o código fosse único, eram exibidos o código e as descrições associadas. Se houvesse códigos múltiplos, estes eram exibidos, bem como as descrições associadas a cada um deles, permitindo a seleção de um código. Se não existisse código, o sistema permitia digitar um novo texto. Se os textos que codificavam (únicos ou múltiplos) já tivessem sido confirmados anteriormente e a freqüência de confirmação tivesse atingido um valor previamente determinado como limite, e houvesse autorização do operador, o código era confirmado automaticamente (aprendizado através do uso), sendo exibidos somente os não confirmados automaticamente. A freqüência para confirmação automática era única para o quesito durante o processamento dos lotes. Se, nesta etapa, algum texto ainda não tivesse sido codificado, era levado para a etapa de resolução de códigos pendentes.

Atribuição de código aos registros individuais

Caso houvesse textos pendentes para encerrar a aplicação de códigos do quesito em questão, o sistema abria uma janela, que permitia realizar este trabalho de forma assistida, apresentando o texto e algumas variáveis auxiliares predefinidas para auxílio da codificação, possibilitando a solução da pendência.

Uma vez codificado o conjunto de textos extraídos do lote, os arquivos originais passavam por um batimento com o arquivo das descrições grupadas já codificadas para atribuição do código aos registros individuais.

Acompanhamento da codificação

O acompanhamento foi realizado por lote de codificação e tema. O lote era uma unidade de trabalho para aplicação de código e, assim sendo, o sistema gerava, para cada um, o quadro de status que permitia saber as fases concluídas para cada quesito e as estatísticas de codificação por quesito (percentuais de codificação automática e assistida). Os codificadores só podiam iniciar o trabalho em um novo lote, quando todas as fases de trabalho no lote atual tivessem sido completadas. Este controle era feito de forma automática pelo sistema.

12.5.1.3 Aplicação de códigos

Como foi visto, a aplicação de um código podia acontecer automaticamente, ou decorrer do trabalho do operador, assistido pela rotina de aplicação de códigos. Esse trabalho normalmente passava pelas seguintes etapas de execução: extração de textos, correção de caracteres, agrupamento por códigos, confirmação de códigos, atribuição de códigos e aplicação assistida. Essas etapas serão tratadas agora, sob o ponto de vista da operação da rotina de aplicação de códigos.

a) Rotina de Aplicação de Códigos

A rotina trabalhava os textos resultantes da digitalização e do reconhecimento ótico – textos originais – de uma só vez e com critérios uniformes. Recebido um lote e escolhido um quesito de codificação, a etapa extração de textos separava, entre todas as pessoas investigadas, os textos originais distintos para, em seguida, submetê-los a um interpretador que, após a utilização de recursos de divisão em palavras, eliminação de preposições, artigos, plurais e gênero, substituição de letras e uso de sinônimos, fazia uma pesquisa, normal e fonética, em um banco descritor temático de textos.

Todos os textos originais que não encontravam correspondência no banco descritor, por insuficiência deste ou por erros de grafia e/ou reconhecimento ótico, eram objeto de trabalho do operador na etapa seguinte, correção de caracteres. Nessa etapa, o operador podia manter ou adequar os textos, para que, depois disso, fossem novamente pesquisados junto ao banco.

Na etapa seguinte, agrupamento por códigos, o sistema classificava, sem o auxílio do operador, cada texto modificado ou não no estágio anterior, em uma das seguintes situações: código único, código múltiplo e sem código. Na seqüência, o operador realizava a confirmação de códigos, onde as sugestões oferecidas pela rotina eram analisadas, podendo ser confirmadas ou levadas, juntamente com os textos sem código, à condição de pendência, para serem resolvidas mais tarde, quase sempre, com o auxílio de variáveis auxiliares.

Passando à etapa atribuição de códigos, a rotina atribuía os códigos aos registros das pessoas que constituíam o lote original, permanecendo ainda sem informação os correspondentes aos textos pendentes, cuja solução acontecia durante o estágio subseqüente, o de aplicação assistida.

Na aplicação assistida - última etapa da aplicação de códigos – a aplicação de código era realizada pessoa a pessoa, e o operador podia, novamente, fazer a adequação dos textos através da correção de caracteres. Dependendo do quesito a ser codificado, era possível também contar com o auxílio de variáveis auxiliares, ou seja, de outros dados originalmente informados no questionário, a saber:

- quesito ocupação – as variáveis auxiliares eram: atividade – quesito 4.46; espécie de curso mais elevado concluído – quesito 4.35; rendimento de trabalho – quesito 4.51 e 4.52; posição na ocupação - quesito 4.47 e número de empregados - quesito 4.49;

- quesito atividade – eram as seguintes as variáveis auxiliares: ocupação – quesito 4.45; rendimento de trabalho – quesito 4.51 e 4.52; posição na ocupação – quesito 4.47 e número de empregados – quesito 4.49;

- quesito curso – as variáveis eram: curso mais elevado que freqüentou – quesito 4.32; última série concluída com aprovação – quesito 4.33; se concluiu o curso no qual estudou – quesito 4.34;

- quesito "município de residência em 31/07/1995" – a variável auxiliar era o quesito 4.26 – UF ou país estrangeiro de residência em 31/07/1995; e

- quesito "UF ou país estrangeiro de residência em 31/07/1995" – a variável auxiliar era o quesito 4.25 – município de residência em 31/07/1995.

Ainda na aplicação assistida, a necessidade de esclarecimento das situações de dúvida de aplicação de códigos podia determinar a consulta aos técnicos das unidades regionais e/ou aos especialistas temáticos da Diretoria de Pesquisas.

A automatização podia ocorrer, após autorização do orientador, nas etapas correção de caracteres e/ou confirmação de códigos, quando um mesmo procedimento se repetia a partir de um número de vezes. Nesse caso, o limite exigido para o início do tratamento automático variou, na correção de caracteres, de 3 a 5 ocorrências, dependendo do quesito; já para confirmação de códigos, foi de 5 ocorrências em Religião, Migração e Curso, enquanto em Ocupação e Atividade estabeleceu-se, inicialmente, 10 ocorrências, baixadas para 5 com o andamento do trabalho.

Especificamente para algumas declarações dos quesitos Ocupação e Atividade, houve a determinação de que os operadores utilizassem o recurso de aplicação de códigos genéricos – códigos alfanuméricos. No quesito Ocupação, fez-se uso dessa alternativa desde o início do processo de aplicação de códigos, enquanto no quesito Atividade a implantação aconteceu com o trabalho já em andamento.

A utilização desse recurso teve como objetivo viabilizar o cumprimento dos prazos para conclusão da fase Aplicação de Códigos. Foi bastante útil nas situações de textos muito freqüentes, aos quais estavam associados códigos múltiplos, cujas atribuições exigiam sempre a passagem pela etapa aplicação assistida, o que aumentava, em muito, o tempo de trabalho.

Para essas situações, os respectivos códigos específicos finais foram atribuídos após terminada a fase Aplicação de Códigos, através de um programa especial, que implementava uma tabela de conversão ou de atribuição de código, estabelecida em conjunto com especialistas temáticos.

Para essa conversão, partia-se do código genérico atribuído e, através da consulta às variáveis auxiliares, fazia-se a decodificação:

- no caso do genérico do quesito ocupação, eram consultadas sempre a posição na ocupação e o código da atividade e, em alguns casos, a variável número de empregados; e

- para o genérico do quesito atividade, utilizava-se, necessariamente, o código da ocupação e, dependendo da situação, também a variável posição na ocupação.

A fase Aplicação de Códigos do lote trabalhado só estava encerrada quando o operador fazia o envio ao sistema do lote já codificado, ocasião em que todas as pendências deveriam estar resolvidas. Em tempo de produção, o sistema podia ser consultado sobre as seguintes estatísticas sobre o lote:

- datas de início e término do trabalho;

- número de registros lidos;

- números de textos extraídos, distintos e corrigidos;

- números de textos com códigos únicos, múltiplos e sem códigos; e

- números de textos distintos levados à aplicação assistida e de pessoas correspondentes

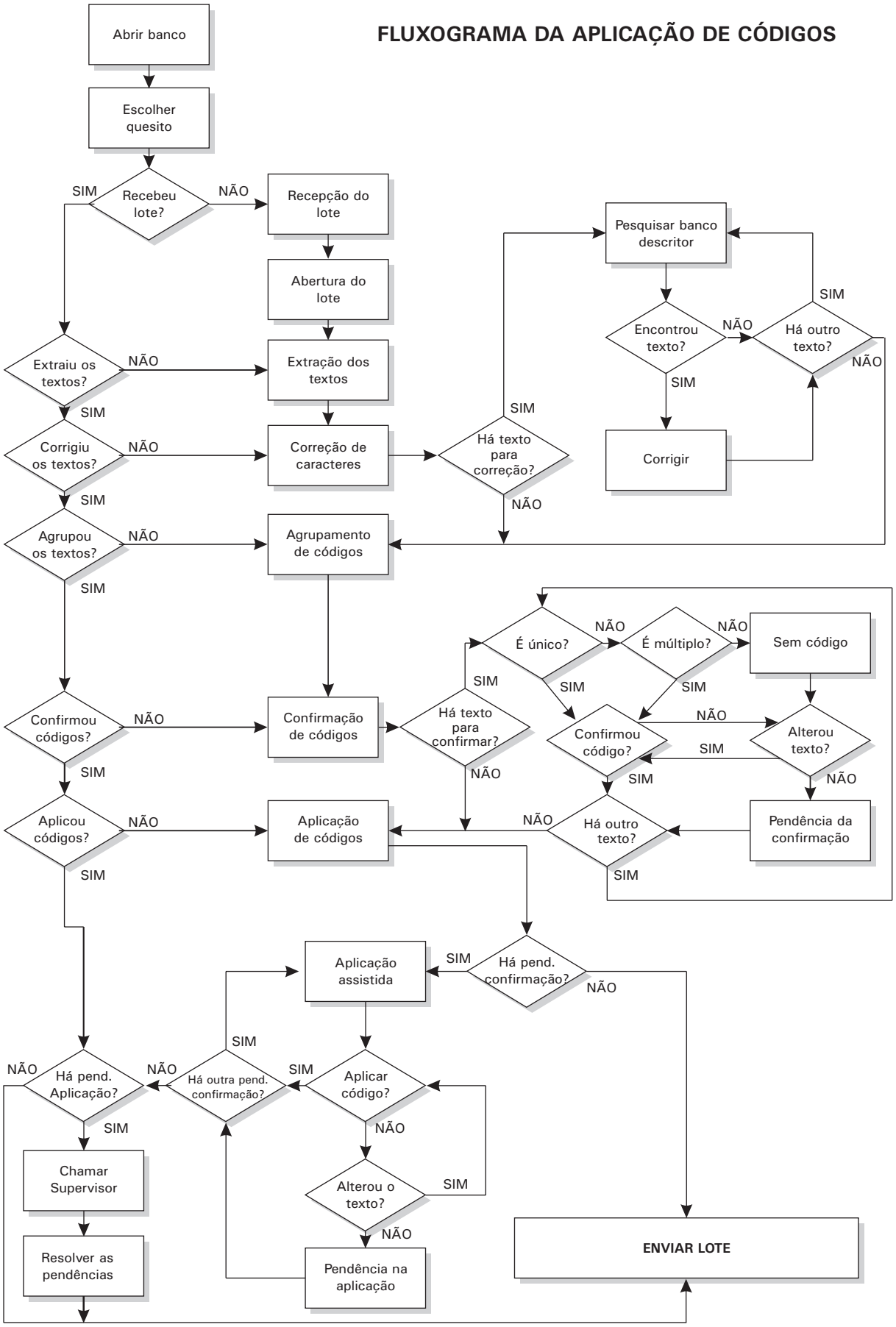
O esquema seguinte ilustra o fluxo de trabalho da fase Aplicação de Códigos, enquanto a Tabela 12.18 apresenta valores para alguns indicadores do processo.

Tabela 12.18 - Alguns indicadores relativos ao processo de aplicação de códigos

Indicadores	Quesito								
	Religião	Unidades da Federação/ País nasc.	Unidades da Federação/ País ant.	Município 5 anos	Unidades da Federação/ País 5 anos	Mun/Unidades da Federação/p. est. e trab.	Curso	Ocupação	Atividade
Tempo médio (em dias) de duração da aplicação de códigos por lote	3,2	1,2	1,3	2,4	1,2	1,2	1,4	33,7	32,8
Média de Registros trabalhados por dia	29 405	20 409	18 660	9 990	20 570	19 495	2 090	1 076	1 094
Média de textos extraídos trabalhados por dia	170	49	50	99	25	75	69	70	112
% textos corrigidos, em relação aos textos extraídos	45,22	55,29	61,23	24,94	65,08	29,14	20,72	24,59	38,43
% textos confirmados, em relação aos textos extraídos	7,39	6,57	4,68	24,00	9,80	50,15	15,05	7,21	4,40
% de textos distintos levados a aplicação assistida, em relação aos textos extraídos	55,32	95,11	96,78	75,01	93,88	83,35	91,74	85,38	59,42
% de registros levados a aplicação assistida, em relação ao total de registros	3,26	3,21	2,89	5,78	1,60	16,37	89,02	50,14	44,82

Fonte: IBGE, Censo Demográfico 2000.

Figura 12.3 – Fluxograma da aplicação de códigos



b) Treinamento e equipes de trabalho

As equipes para o trabalho de aplicação de códigos eram formadas por operadores/codificadores e por orientadores. Os operadores eram técnicos temporários, contratados para as atividades de apuração do censo, enquanto os orientadores faziam parte do quadro permanente do IBGE. Cada equipe desenvolvia a aplicação de códigos em um determinado tema. Além das tarefas de acompanhamento e supervisão dos trabalhos, cabia também aos orientadores avaliar as situações de automatização da correção de caracteres e da confirmação de códigos, cuja validação dependia de sua autorização, mediante o uso de senha específica.

A aplicação de códigos foi realizada no Rio de Janeiro, em duas instalações do IBGE e em dois turnos de trabalho, com exceção do tema Religião, cuja equipe atuou em regime de turno único. Especificamente no quesito Atividade, só houve um segundo turno durante os últimos três meses de trabalho. A tabela seguinte apresenta os quantitativos das equipes e os tempos gastos no trabalho, segundo cada quesito de codificação.

Tabela 12.19 - Tamanhos das equipes de aplicação de códigos e períodos de realização do trabalho, segundo os temas de codificação

Temas de codificação	Codificadores	Orientadores	Data	
			Inicial	Final
Total	**102**	**17**		
Religião	5	2	03/08/01	21/11/01
Curso	10	2	01/08/01	01/11/01
Migração	10	2	13/08/01	09/11/01
Ocupação	34	4	13/08/01	07/02/02
Atividade	43	7	18/07/01	11/03/02

Fonte: IBGE, Censo Demográfico 2000.

Os codificadores foram contratados em maio de 2001, com uma antecedência média de dois meses do início efetivo da aplicação de códigos Esse prazo atendeu às peculiaridades do treinamento das equipes, que foi realizado, inicialmente, para duas turmas, em duas etapas. A primeira delas, com a duração de 5 dias, abrangia alguns aspectos teóricos do censo e o conhecimento da rotina de aplicação de códigos, tendo como base o manual Procedimentos Operacionais para o Sistema de Codificação. A segunda etapa de capacitação, que durou em média 60 dias, estava totalmente voltada para a atividade prática do trabalho, tendo como ênfase qualificar os operadores no conhecimento dos bancos descritores e dar-lhes agilidade no uso da rotina; para tanto, durante esse período, foi realizada a aplicação de códigos, utilizando-se as declarações obtidas nos questionários do Censo Experimental.

Posteriormente, uma nova turma de operadores foi treinada, visando à criação do segundo turno de trabalho para aplicação de códigos no tema Atividade.

12.5.1.4 Verificação de códigos

A Verificação de Códigos teve como objetivo contribuir para a melhoria da qualidade do trabalho da fase Aplicação de Códigos. Operacionalmente, a tarefa consistiu em fazer, por amostragem, uma segunda atribuição de códigos

para todos os lotes de cada quesito de codificação, visando identificar as situações de divergência entre os códigos aplicados pelos dois operadores, codificador e verificador.

O trabalho de verificação foi realizado por 15 técnicos que, diferentemente da fase Aplicação de Códigos, atuavam em qualquer dos nove quesitos de codificação. A tarefa tinha início após a liberação do lote de codificação e da constituição dos correspondentes arquivos intermediários. Uma rotina especial fazia a gravação desses arquivos, um para cada lote, cujos registros continham a identificação da pessoa, o texto original, o texto após o estágio correção de caracteres, o código atribuído e, dependendo do quesito, as variáveis auxiliares. Constava, também, desse arquivo uma classificação que mostrava como a rotina e/ou o(s) operador(es) tratou(aram) - indicando se houve ou não procedimento automático – o texto de cada pessoa, em cada estágio do trabalho da Aplicação de Códigos.

a) Rotina de Verificação de códigos

Para a atividade de verificação de códigos, desenvolveu-se um aplicativo em cuja operação acontecia, em linhas gerais, o seguinte:

- liberado um lote de codificação, através do seu arquivo intermediário, era constituído o correspondente lote de verificação, que continha somente os registros das pessoas em que, pelo menos, numa das etapas, correção de caracteres e confirmação de códigos, houve a necessidade de intervenção do operador;

- no lote de verificação, era selecionada uma amostra de registros onde o verificador, utilizando a mesma rotina de trabalho do operador, com exceção dos mecanismos de correção e confirmação automáticos, fazia a nova aplicação de códigos;

- para cada registro da amostra, após o trabalho de verificação, eram acrescentadas as informações do arquivo intermediário, o texto após a correção de caracteres e o código aplicado pelo verificador;

- eram identificadas as situações de divergência entre os códigos aplicados pelo codificador e verificador, e calculada sua proporção na amostra;

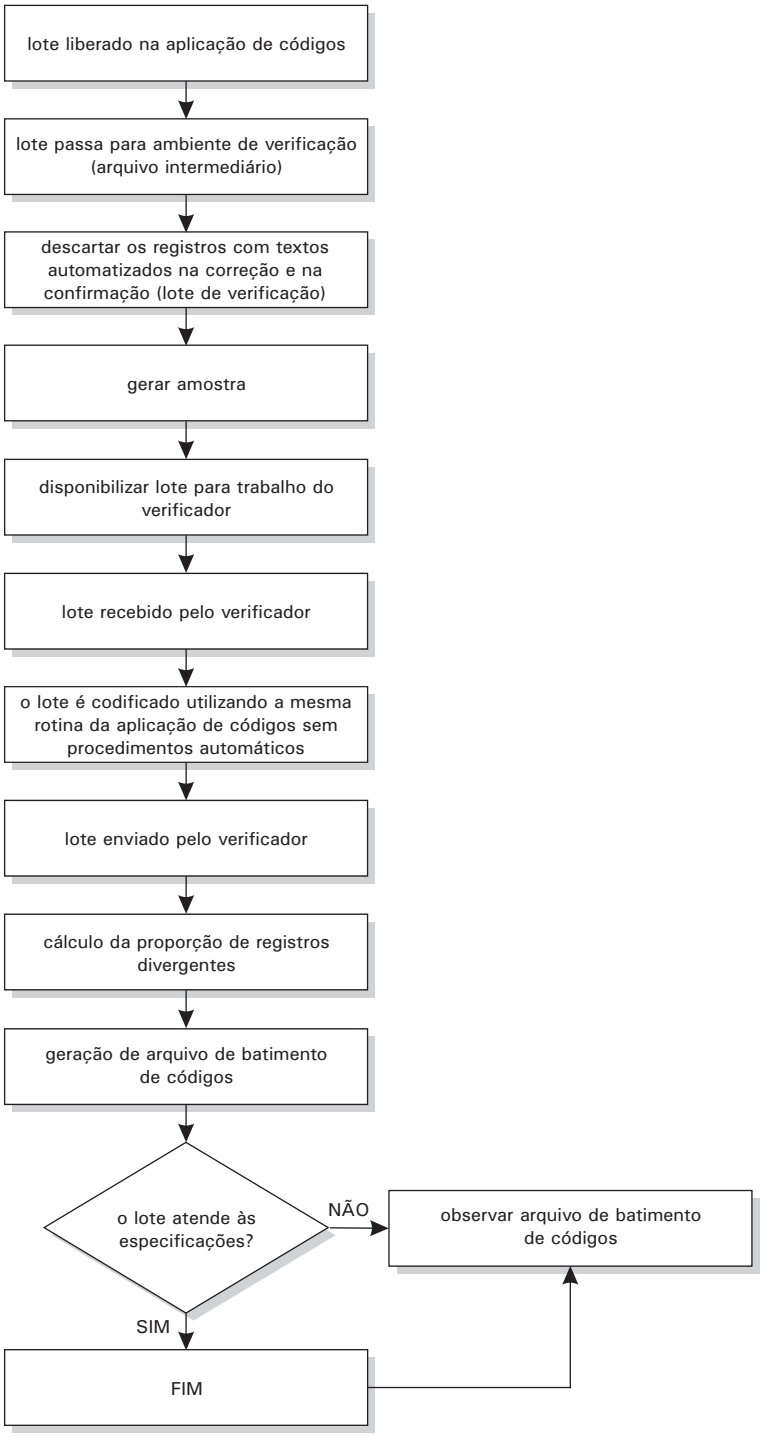
- constituía-se o arquivo de batimento, que continha, apenas, os registros onde ocorreram divergências de códigos, cujo conteúdo apresentava as mesmas informações do arquivo intermediário, pós trabalho de verificação; e

- dependendo do valor da proporção de códigos divergentes, era impresso, para investigação, o arquivo de batimento do lote de verificação.

O fluxo de trabalho, na etapa Verificação de Códigos, é apresentado a seguir.

Figura 12.4 – Fluxograma da verificação de códigos

Fluxograma da etapa de verificação


lote liberado na aplicação de códigos
lote passa para ambiente de verificação (arquivo intermediário)
descartar os registros com textos automatizados na correção e na confirmação (lote de verificação)
gerar amostra
disponibilizar lote para trabalho do verificador
lote recebido pelo verificador
o lote é codificado utilizando a mesma rotina da aplicação de códigos sem procedimentos automáticos
lote enviado pelo verificador
cálculo da proporção de registros divergentes
geração de arquivo de batimento de códigos
o lote atende às especificações?
NÃO
observar arquivo de batimento de códigos
SIM
FIM

Por fim, o programa fazia a atualização dos seguintes relatórios de acompanhamento e avaliação da etapa Verificação de Códigos:

- gerenciamento de lotes 1 – apresenta o lote, segundo o quesito de verificação e o estágio de trabalho (codificado, disponível para verificação, amostrado, em trabalho, enviado e calculado);

- gerenciamento de lotes 2 – identifica o verificador, quesito de verificação e as datas de início e término do trabalho do lote;

- quantitativo de lotes – mostra a quantidade de lotes, segundo o estágio de trabalho e o quesito de verificação; e

- tamanhos dos lotes e amostras, segundo o quesito de verificação;

- proporção de códigos divergentes no lote; e

- relatório de lotes não classificados.

b) Critérios e Procedimentos

Levando-se em conta amostragem aleatória simples, a proporção Pvi de códigos divergentes em um determinado lote de verificação i foi estimada por

$$p_{vi} = \frac{d_{vi}}{n_i}$$

onde :

p_{vi} = proporção de códigos divergentes na amostra do lote i;

d_{vi} = número de códigos divergentes na amostra do lote i; e

n_i = número de códigos aplicados na amostra do lote i.

Definiu-se como "classificado", o lote de verificação cuja proporção máxima estimada de registros com códigos divergentes fosse da ordem de 5%. Ainda na amostragem aleatória simples sem reposição, para a estimação da proporção de 3% com um coeficiente de variação – CV – de 25% e o grau de confiança de 99%, tem-se um limite superior para o intervalo de aceitação de 5,25%, o que atendia, aproximadamente, ao limite estabelecido para julgamento do lote.

A especificação de um valor para a proporção de registros com códigos divergentes – ao redor de 5% – como parâmetro, que classificava um lote de verificação, foi feita arbitrariamente. Já a escolha dos valores de p (3%) e CV (25%), obedeceu à disponibilidade de pessoal e à relação entre estimativas dos tempos médios de trabalho dos operadores e verificadores. Mesmo nesse último caso, um certo grau de arbitrariedade aconteceu, pois seria possível outras combinações dos parâmetros p e CV que resultassem em tamanhos de amostra também adequados.

Para todos os lotes identificados como "não classificados", procedeu-se à investigação dos respectivos arquivos de batimento, objetivando esclarecer o que acarretou as divergências entre os códigos aplicados: se erro do operador e/ou do verificador. Assim, a(s) fonte(s) de erros identificada(s), associada(s) a outros critérios de acompanhamento do trabalho, determinavam quais as ações - alertas, retreinamento e implantação de novos procedimentos – seriam efetivadas junto aos operadores, orientadores e verificadores, tendo em vista melhorar a qualidade do trabalho.

Como foi visto, os pvi estimam, para todos os quesitos, as proporções de códigos divergentes nos lotes de verificação. No entanto, é necessário também fornecer uma indicação do resultado do trabalho de aplicação de códigos, tendo como referência os correspondentes lotes de codificação. Ou seja, devemos levar em conta para o cálculo da nova medida, a parcela dos códigos aplicados através de mecanismos automáticos de correção de caracteres e confirmação de códigos, simultaneamente.

Quando se toma somente os registros das pessoas em cujos códigos houve, necessariamente, a participação do operador, ou seja, sem o apoio de mecanismos automáticos – lotes de verificação – tem-se que a proporção de registros com códigos divergentes no respectivo lote de codificação será, no máximo, igual a estimada para o lote de verificação.

Assim, tendo em conta que na Aplicação de Códigos, os textos só são passíveis de tratamento automático após as situações superarem determinados limites de freqüência e serem avaliadas pelos orientadores, os códigos decorrentes dos procedimentos simultâneos de correção de textos e confirmação de códigos podem ser considerados corretos. Com base nessa hipótese, a proporção de códigos divergentes estimada para o lote de codificação i será

$$P_{ci} = p_{vi} \cdot N_i / M_i$$

onde N_i e M_i são, respectivamente, o tamanho do lote de verificação e de codificação, ou seja, o número de códigos obtidos sem mecanismos automáticos simultâneos e o número total de códigos atribuídos no lote.

A seguir, são apresentados gráficos para cada quesito de codificação que mostram as freqüências relativas – fr – dos lotes de codificação, segundo as proporções estimadas de códigos divergentes, relativas aos 215 lotes trabalhados. Deve-se registrar que essas proporções – eixo horizontal do gráfico – não representam equívocos de aplicação de códigos, mas tão somente as proporções estimadas de não coincidência entre códigos aplicados pelo operadores e verificadores.

Gráfico 12.5 - Quesito Religião
Freqüências relativas - fr - dos lotes de codificação, segundo as proporções estimadas de códigos divergentes

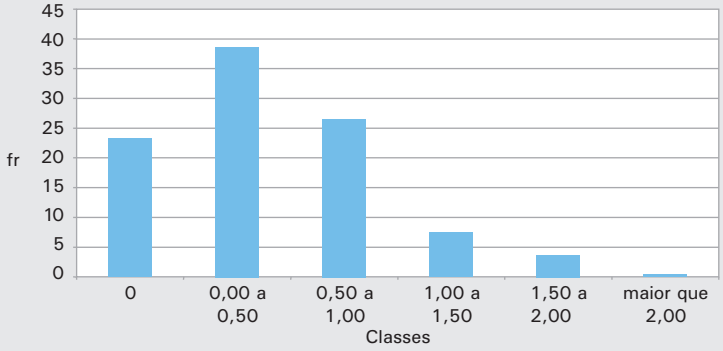


Gráfico 12.6 - Quesito Curso

Freqüências relativas - fr - dos lotes de codificação, segundo as proporções estimadas de códigos divergentes

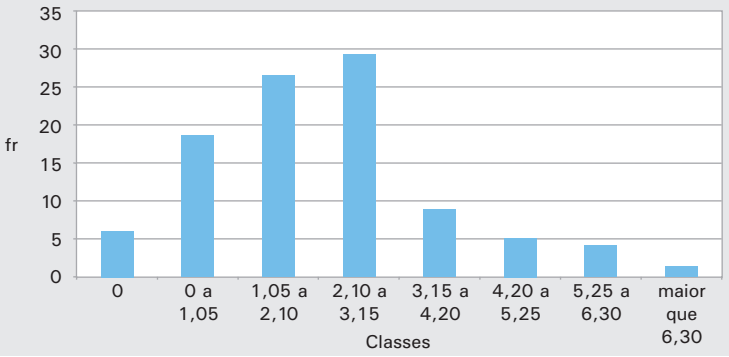


Gráfico 12.7 - Tema Migração - Quesito 4.21

Freqüências relativas - fr - dos lotes de codificação, segundo as proporções estimadas de códigos divergentes

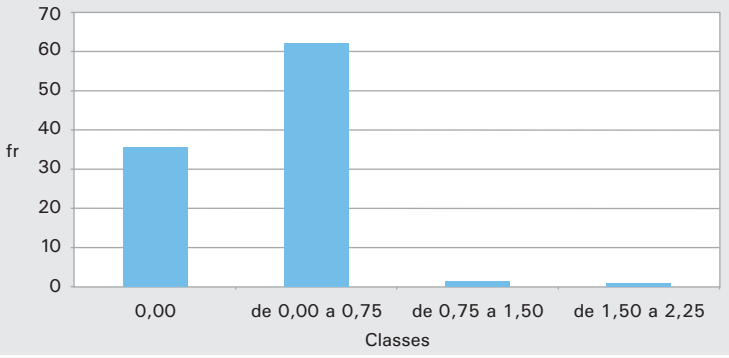


Gráfico 12.8 - Freqüências relativas - fr - dos lotes de codificação, segundo as proporções estimadas de códigos divergentes

Tema Migração - Quesito 4.23

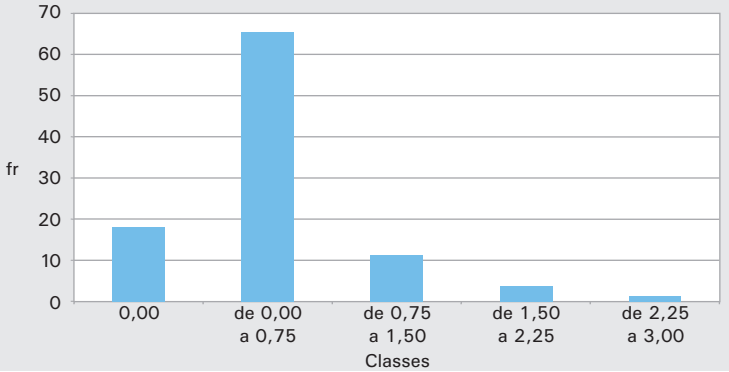


Gráfico 12.9 - Freqüências relativas - fr - dos lotes de codificação, segundo as proporções estimadas de códigos divergentes

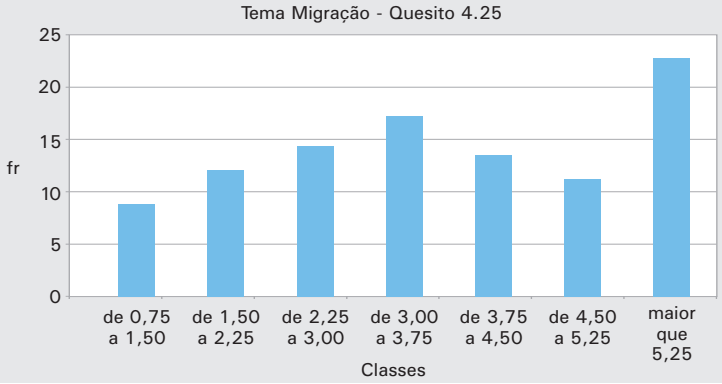


Gráfico 12.10 - Freqüências relativas - fr - dos lotes de codificação, segundo as proporções estimadas de códigos divergentes

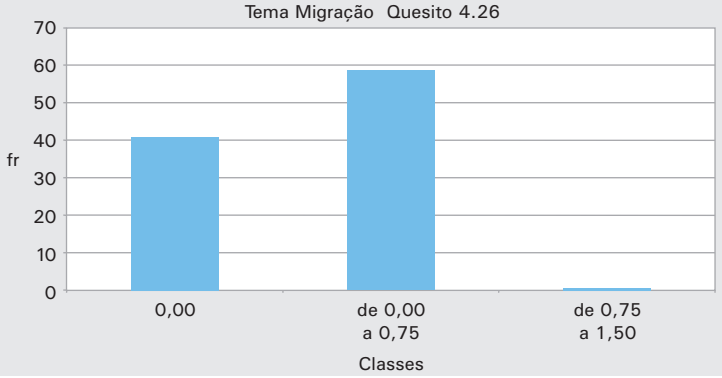


Gráfico 12.11 - Freqüências relativas - fr - dos lotes de codificação, segundo as proporções estimadas de códigos divergentes

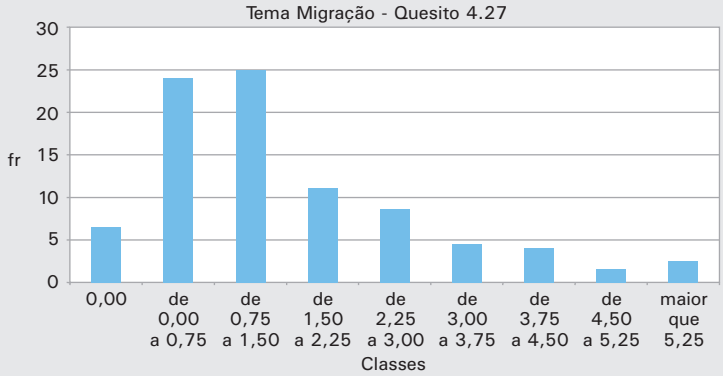


Gráfico 12.12 - Freqüências relativas - fr - dos lotes de codificação, segundo as proporções estimadas de códigos divergentes

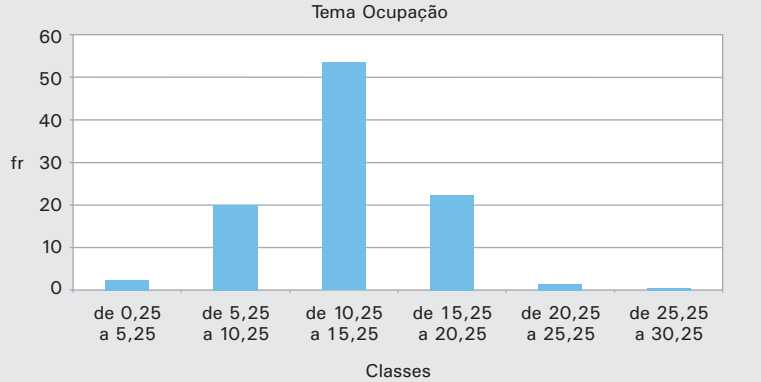
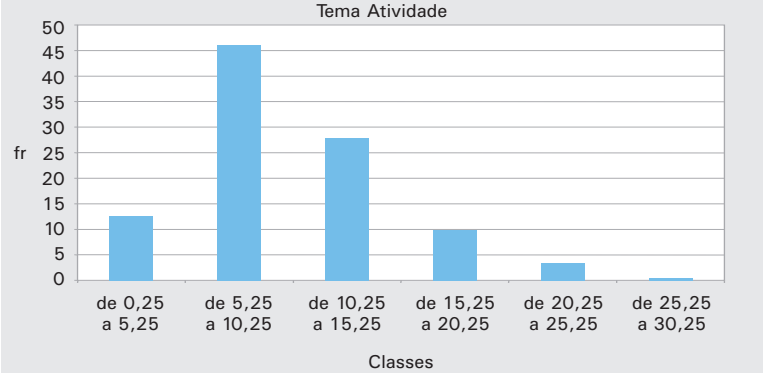


Gráfico 12.13 - Freqüências relativas - fr - dos lotes de codificação, segundo as proporções estimadas de códigos divergentes



12.5.2 Crítica e imputação dos dados

O planejamento dos trabalhos de apuração dos dados do Questionário da Amostra do Censo Demográfico 2000, com vistas à execução da crítica de incompatibilidades, partiu do princípio de que o sistema de detecção e imputação dos erros seria o DIA, ou seja, o mesmo utilizado no Censo Demográfico de 1991. Naquela oportunidade, as críticas de incompatibilidades foram, em quase sua totalidade, desenvolvidas de forma centralizada e apenas um conjunto mínimo, o correspondente às críticas entre registros, foi corrigido de forma manual, descentralizadamente, em pólos de apuração estruturados em vinte Unidades da Federação.

O objetivo traçado para o Censo 2000 era tratar toda a crítica dos dados do CD 1.02 de maneira centralizada, buscando a eliminação total do processo manual de correção, executado no Censo de 1991, mesmo sabendo das dificuldades do DIA em trabalhar com regras de crítica entre registros. O ponto de partida para consecução dessa meta foram os testes com os arquivos das informações da Segunda Prova Piloto.

Assim, inicialmente, foi escrito o conjunto de regras de crítica, descrevendo as inconsistências relativas às variáveis envolvidas nas críticas entre registros: sexo, relação com a pessoa responsável pelo domicílio, relação com a pessoa responsável pela família, número da família, idade do entrevistado, natureza da união e estado civil.

Durante esse processo de elaboração do conjunto de críticas, tendo em vista contornar a dificuldade do sistema DIA em trabalhar com regras de crítica entre registros, houve a necessidade de criação de um número razoável de variáveis auxiliares, a partir das variáveis originais do questionário.

Terminado o trabalho de elaboração das regras de crítica entre registros, verificou-se que o sistema DIA não obtinha êxito no procedimento de geração do conjunto completo de regras de crítica. Após várias tentativas, alternando-se as estratégias de criação das regras de incompatibilidades, permanecendo essa dificuldade, decidiu-se pelo abandono do sistema DIA para tratar as críticas entre registros.

Todo o desenvolvimento desse trabalho pode ser melhor conhecido, consultando-se o texto *Relatório sobre a experiência de tratamento automático de crítica entre registros, com vistas ao censo demográfico do ano 2000* (1999).

O abandono do sistema DIA determinou estudos para a busca de outra alternativa para a crítica entre registros dos dados do Questionário da Amostra. Os resultados desses estudos apontaram para duas vertentes metodológicas. A primeira, através de procedimentos automáticos de imputação, utilizando o sistema New Imputation Methodology – NIM para corrigir as inconsistências, considerando-se apenas os domicílios com até 8 moradores; a segunda, através do sistema Integrated Microcomputer Processing System – IMPS, atuando nos domicílios com 9 até 38 moradores, a partir de imputação manual.

O sistema NIM executa a correção dos dados, a partir da obtenção de domicílios doadores, selecionados dos estoques de domicílios sem nenhuma inconsistência nas informações, segundo o número de moradores. Nesse caso, foi necessária a formação de oito estratos de depuração.

Já o IMPS é um sistema de detecção de erros e, por essa razão, a correção deve ser realizada manualmente, em tela de microcomputador, através da análise das mensagens de erro. Técnicos treinados estiveram encarregados da execução desse trabalho, sendo os domicílios agrupados em um único estrato.

Ainda com relação a crítica entre registros, é necessário dizer que os cortes estabelecidos - 8 e 38 moradores – não se deram de forma arbitrária. Estabeleceu-se o critério, tendo em vista a preservação, a partir dos estoques de domicílios doadores , das características originais dos domicílios a serem corrigidos.

O critério de corte, domicílios com até oito moradores, levou em conta os resultados dos testes realizados com os arquivos da Segunda Prova Piloto. Em princípio, pretendia-se utilizar unicamente o NIM para tratar as críticas entre registros, envolvendo todos os domicílios. No entanto, durante esses testes, verificou-se que os resultados da imputação para domicílios com 12 e 16 moradores não foram satisfatórios. Aconteceu que, após a imputação, os domicílios viriam a ter suas características originais bastante alteradas. Houve, também, uma avaliação para os domicílios com 9 e 10 moradores, sendo considerados insuficientes os estoques de domicílios doadores para essas dimensões.

De acordo com o planejamento dos trabalhos para a crítica através do NIM, caso, em algum estrato, não se conseguisse um domicílio doador, o domicílio com erro era tratado pela alternativa IMPS.

12.5.2.1 Crítica entre registros

O objetivo deste item é descrever o procedimento utilizado para a crítica e imputação dos dados investigados, no Questionário da Amostra, sob a ótica da comparação entre registros de um mesmo domicílio. Além disso, são apresenta-

dos alguns resultados de uma exploração inicial feita nos arquivos de registro de execução dessa atividade, que indicam os tipos de análise que podem ser realizadas sobre o processo. Essa análise inicial oferece uma idéia da relevância das informações disponíveis para uma avaliação mais detalhada que tenha por objetivo entender o que aconteceu no Censo 2000 em termos de imputação, e ajudar no planejamento do próximo censo demográfico, ou de pesquisas correlatas.

Como já descrito em itens anteriores, na etapa de crítica de consistência do Censo Demográfico 2000, aproveitando a experiência adquirida no Censo de 1991, foi definida a utilização do programa DIA (Detección e Imputación Automática). O DIA trabalha muito bem no que se refere à consistência de variáveis dentro do mesmo registro (seja de domicílios ou pessoas), mas tem a limitação de não possuir ferramentas para verificar as regras entre registros (isto é, entre pessoas distintas, ou entre uma variável de domicílio e variáveis de pessoas).

A utilização do sistema DIA na crítica entre registros do Conjunto Universo foi possível em função do pequeno número de variáveis a serem criticadas e da estratégia adotada, que foi a de construir um único registro com as informações de todas as pessoas que seriam objeto dessa crítica. Esse procedimento está descrito no item 12.3.2.4, que trata da apuração do Conjunto Universo.

Para enfrentar este problema, buscou-se o NIM – Nearest-neighbour Imputation Methodology, do Statistics Canada, na época somente um protótipo, cujo nome era New Imputation Methodology, mas que apresentava uma série de vantagens técnicas e operacionais, a saber:

- a imputação é totalmente automática, bastando definir as regras de consistência (isto é, não é necessário definir também as regras de imputação);
- a imputação é baseada em um único doador por questionário falhado (domicílio);
- segue a filosofia de Feleggi e Holt, no sentido de alterar o menor número de variáveis do questionário;
- trabalha com arquivos comuns de tipo texto; e,
- o software foi cedido ao IBGE "em aberto", isto é, com os programas-fonte.

Por outro lado, devido à própria característica de imputação dos domicílios com erro a partir de doadores (domicílios "bons"), essa imputação baseia-se na busca de um doador que tenha a menor distância, no sentido de maior semelhança do conteúdo das variáveis. Por esse motivo, o NIM trabalha separando os domicílios por estratos de número de pessoas (domicílios com 1 pessoa, domicílios com 2 pessoas, domicílios com 3 pessoas, ..., domicílios com n pessoas), posto que este procedimento padroniza a utilização de doadores de mesmo tamanho (mesmo número de pessoas) para o domicílio com erro.

Ao processar um estrato qualquer, é de se esperar que este tenha um número mínimo de doadores tal que possa assegurar uma imputação de boa qualidade (distâncias pequenas, não utilização de mesmo doador, etc.). Isto não é possível de se garantir para os estratos com um maior número de pessoas (domicílios com 12 pessoas, por exemplo). Para estratos maiores, é possível não conseguir doador, impossibilitando o uso do método.

Nesse sentido, limitou-se a oito o número de estratos a serem trabalhados pelo NIM (desde domicílios com 1 pessoa, até domicílios com 8 pessoas). Os domicílios com mais de 8 pessoas foram tratados por um outro método, o IMPS

(Integrated Microcomputer Processing System), fornecido pelo Census Bureau, dos Estados Unidos. O IMPS necessita que sejam estipuladas tanto as regras de consistência como as de imputação, ou que a imputação seja operada manualmente por pessoal qualificado (os questionários com erro são apresentados na tela com seus respectivos erros, e os operadores definem ações de imputação para "limpá-los").

O ideal seria que todas as imputações pudessem ser definidas de maneira automática, sem intervenção humana, mas isto não foi de todo possível, dada a complexidade das relações de estrutura do questionário (relações entre pessoas). Isto obrigou que se dispusesse de um conjunto de operadores que processassem os questionários com mais de 8 pessoas.

Descrição dos procedimentos

Resumidamente e, em termos lógicos, o processo consistia na formatação do arquivo para extração das variáveis que seriam trabalhadas (somente aquelas envolvidas com a estrutura do domicílio). Em seguida, este arquivo era separado em 9 arquivos, um para cada estrato a ser tratado pelo NIM, e um único estrato com os domicílios com mais de 8 pessoas, a ser tratado pelo IMPS. Esses procedimentos foram realizados independentemente em cada um dos 67 lotes de questionários formados especificamente para essa etapa de crítica entre registros. Os lotes foram definidos, considerando os seguintes critérios:

- possuir, pelo menos, 1000 questionários em cada estrato, com o objetivo de formar massa crítica suficiente para a alocação de registros doadores similares aos que falhassem em alguma regra de crítica. esse valor foi definido com base nos testes de uso do NIM realizados pelo Statistics Canada. Por outro lado, aumentar a quantidade mínima de questionários implicaria a necessidade de agregação de Unidades da Federação menos populosas em um único lote. Também não seria possível, na maioria dos casos, obter pelo menos um lote de domicílios com situação "rural" por Unidade da Federação;

- considerar a separação por Unidade da Federação, desde que o primeiro critério continuasse válido;

- considerar a separação por situação do domicílio (urbano ou rural), desde que também o primeiro critério continuasse válido;

- ter um limite máximo de registros para um estrato (aproximadamente 20 000), de acordo com a capacidade e tempo de processamento do sistema, porém, não foi definido o tamanho máximo para um lote;

- obter o menor número possível de lotes, porém respeitando os critérios propostos.

A composição final dos lotes usados para a execução dessa etapa de apuração está apresentado no Quadro 12.3.

Quadro 12.3 - Composição dos lotes de questionários para a realização da etapa de crítica e imputação entre registros e em número de domicílios particulares ocupados na amostra

Unidades da Federação	Lote	Mesorregião	Número de domicílios particulares ocupados na amostra
11	1	lote único	42 966
12, 14 e 16	1	lote único	38 027
13	1	lote único	63 694
15	1	Urbana	96 317
	2	Rural	48 789
17	1	lote único	42 648
21	1	Urbana	85 656
	2	Rural	64 042
22	1	Urbana	54 851
	2	Rural	39 344
23	1	2 e 3 (urbana)	82 445
	2	1, 4 , 5, 6 , 7 (urbana)	56 852
	3	Rural	60 877
24	1	lote único	92 118
25	1	Urbana	77 603
	2	Rural	39 292
26	1	1, 2, 4 (urbana)	45 488
	2	3 (urbana)	36 537
	3	5 (urbana)	84 741
	4	Rural	57 226
27	1	Urbana	51 925
	2	Rural	25 591
28	1	lote único	54 825
29	1	1, 2, 3 (urbana)	55 675
	2	4, 5 (urbana)	111 854
	3	6, 7 (urbana)	76 689
	4	1 a 5 (rural)	75 205
	5	6 e 7 (rural)	57 054
31	1	1, 2, 3, 4 (urbana)	57 329
	2	5, 9 (urbana)	77 556
	3	6, 8, 10 (urbana)	34 533
	4	7 (exceto município 6200) (urbana)	86 121
	5	7 (município 6200 (BH)) (urbana)	63 308
	6	11, 12 (urbana)	71 208
	7	1, 2, 3, 4, 5, 6, 9, 10 (rural)	82 207
	8	7, 8, 11, 12 (rural)	40 561
32	1	lote único	98 067
33	1	1, 2, 3, 4, 5 (urbana e rural)	89 926
	2	6 (micros 14, 15, 16, 17,18 - município 456 a 3302) (urbana)	91 511
	3	6 (micro 18 – município 4557)	180 173
	4	6 (micro 18 – demais municípios – urbana) e rural meso 6	75 899
35	1	1, 8 (urbana)	76 351
	2	2, 9, 14 (urbana)	85 102
	3	3, 4, 5 (urbana)	81 757
	4	6, 13 (urbana)	86 205
	5	7 (urbana)	88 892
	6	10, 11, 12 (urbana)	127 101
	7	15 (micros 57, 58, 59 -urbana)	83 139
	8	15 (micros 60, 62, 63) (urbana)	87 806
	9	15 (micro 61 - exceto município 50308) (urbana)	65 925
	10	15 (micro 61 (município 50308 - áreas 1, 2, 3)	84 902
	11	15 (micro 61 - município 50308 - áreas 4, 5) (urbana)	93 459
	12	15 (micro 61 - município 50308 - áreas 6, 7, 8) (urbana)	104 579
	13	Rural	88 436
41	1	1, 2, 3, 4, 6 (urbana)	134 828
	2	5, 7, 8, 9, 10 (urbana)	122 985
	3	Rural	74 959
42	1	1, 2, 3 (urbana e rural)	92 233
	2	4, 5, 6 (urbana e rural)	99 508
43	1	1, 6 (urbana e rural)	107 017
	2	2, 3, 4, 7 (urbana e rural)	108 293
	3	5 (urbana e rural)	146 080
50	1	lote único	68 316
51	1	lote único	85 402
52	1	1, 2, 4, 5 (urbana e rural)	90 590
	2	3 (urbana e rural)	82 866
53	1	lote único	53 940

Fonte: IBGE, Censo Demográfico 2000, Sistema de Indicadores Gerenciais da Coleta.

Os estratos de 1 a 8 pessoas eram processados pelo NIM, que produzia como saída um arquivo imputado, de mesmo formato que o arquivo de entrada e um arquivo de controle ou ocorrências, aonde eram registradas, para cada domicílio, as seguintes informações (entre outras):

- regras falhadas, com a identificação da regra;
- variáveis imputadas, com a identificação da variável, o valor antigo e o valor novo.

O estrato de 9 e mais pessoas era tratado pelo IMPS, cujo processo também tinha como saída dois arquivos, o arquivo imputado e o arquivo de controle. Por possuir um componente manual (operado por pessoas), seu arquivo de controle não continha os registros de variáveis imputadas, somente o de regras falhadas.

Estes arquivos de controle são os arquivos usados para a exploração aqui descrita. É importante ressaltar que os arquivos de controle de imputação, recolhidos no processo do IMPS, não contêm informações sobre as variáveis imputadas, o que é uma limitação nas análises de imputação. Isso significa que, na maioria dos casos, o universo de análise é o dos domicílios com até 8 pessoas. Também não se encontram nestes arquivos as informações sobre a imputação dos domicílios em outros processos posteriores, como o DIA e a imputação dos rendimentos.

Tratamento das omissões da variável "espécie do domicílio"

Embora a variável "espécie do domicílio" estivesse relacionada, mais diretamente, ao Aplicativo Domicílio, fez parte também do conjunto de tabelas de decisão lógica do NIM. Em algumas dessas tabelas, a crítica relacionava a categoria da variável "espécie do domicílio" com as variáveis "relação com a pessoa responsável pelo domicílio" e "número da família. Por esse motivo, era preciso que houvesse informação para essa variável.

Para corrigir as omissões de informação da variável "espécie do domicílio", foi utilizada a mesma estratégia definida para os dados do Conjunto Universo, descrita no item 12.3.2.4. A omissão era detectada e corrigida durante o processo de formação dos estratos a serem submetidos ao NIM.

As variáveis envolvidas na estrutura do domicílio e que foram objeto dessa etapa de crítica entre registro são:

Quadro 12.4 - Variáveis envolvidas no tratamento das omissões

V0201	Espécie do domicílio (variável do domicílio, somente para consulta, nunca imputada)
V0401	Sexo da pessoa
V0402	Relação com o responsável pelo domicílio
V0403	Relação com o responsável pela família
V0404	Número da família
V0436	Vive em companhia de cônjuge
V0437	Natureza da última união
V0438	Estado civil
V4007	Faixa de idade
V4667	Indicador de fecundidade

Para dar uma idéia do tipo de crítica que foi realizada, envolvendo as variáveis acima, seguem exemplos descritivos das relações verificadas:

- duas pessoas classificadas como responsável pelo domicílio, V0402, (ou pela família, V0403) e cônjuge do responsável pelo domicílio (ou da família) devem ter informações iguais para a variável estado civil (V0438);
- duas pessoas classificadas como responsável pelo domicílio (ou pela família) e cônjuge do responsável pelo domicílio (ou da família) devem ter informações diferentes para a variável sexo (V0401);
- pessoa menor de 10 anos de idade não pode ter informações nas variáveis que investigam as características de nupcialidade (V0435, V0437 e V0438);
- o número de famílias em um domicílio deve ser igual ao número de pessoas classificadas como responsáveis pela famílias;
- todo domicílio tem que ter apenas um responsável pelo domicílio;
- nenhuma variável pode estar em branco.

Exploração de resultados

Pode-se dizer que, para o total Brasil, 28% dos domicílios tiveram, pelo menos, um erro na consistência de sua estrutura, seja este erro apontado pelo NIM ou IMPS[4]. A Tabela 12.20 traz estes percentuais de erro por UF, os quais também são mostrados no Gráfico 12.14 a seguir. Neste, pode-se ver que a UF com maior percentual de erros foi o Amazonas, com 40%, e a menor, o Rio Grande do Sul, com cerca de 19%.

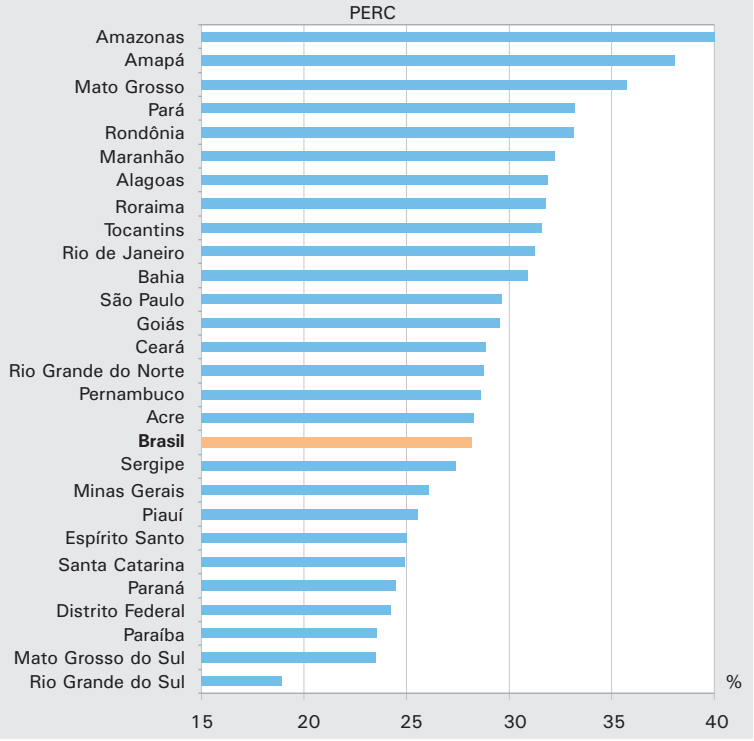
Tabela 12.20 - Número total de domicílios e de domicílios com erro, segundo as Unidades da Federação

Unidades da Federação	Número total de domicílios	Número de domicílios com erro	Percentual de domicílios com erro
Brasil	**5 304 711**	**1 494 397**	**28,17**
Rondônia	43 293	14 350	33,15
Acre	16 818	4 757	28,29
Amazonas	63 970	25 623	40,05
Roraima	9 857	3 133	31,78
Pará	145 992	48 514	33,23
Amapá	11 821	4 502	38,08
Tocantins	43 043	13 599	31,59
Maranhão	150 441	48 536	32,26
Piauí	94 534	24 156	25,55
Ceará	201 143	58 078	28,87
Rio Grande do Norte	92 673	26 669	28,78
Paraíba	117 577	27 698	23,56
Pernambuco	225 649	64 559	28,61
Alagoas	77 896	24 848	31,9
Sergipe	55 161	15 114	27,4
Bahia	378 907	117 216	30,94
Minas Gerais	615 101	160 622	26,11
Espírito Santo	98 820	24 716	25,01
Rio de Janeiro	442 976	138 408	31,25
São Paulo	1 137 154	337 148	29,65
Paraná	336 151	82 235	24,46
Santa Catarina	193 633	48 261	24,92
Rio Grande do Sul	365 827	69 217	18,92
Mato Grosso do Sul	69 401	16 332	23,53
Mato Grosso	86 946	31 053	35,72
Goiás	175 132	51 777	29,56
Distrito Federal	54 795	13 276	24,23

Fonte: IBGE, Censo Demográfico 2000.

[4] Aqui não foram computados os possíveis erros encontrados posteriormente pelos outros processos de consistência (DIA, imputação dos rendimentos, etc.).

Gráfico 12.14 - Percentual de domicílios com erro, por Unidade da Federação



A Tabela 12.21 mostra os totais de erros, segundo os estratos (número de pessoas por domicílio), também mostrados no Gráfico 12.15 abaixo. Os valores variam desde cerca de 13% para o estrato 1 até cerca de 54% para o estrato de 9 ou mais pessoas (processado pelo IMPS).

Tabela 12.21 - Número total de domicílios e de domicílios com erro, segundo os estratos

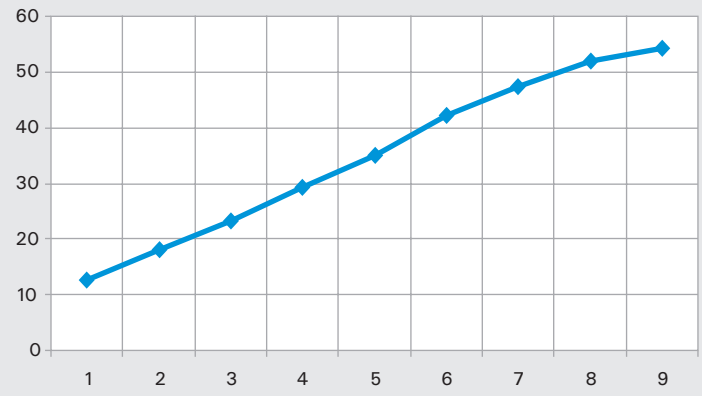
Estrato	Número total de domicílios	Número de domicílios com erro	Percentual de domicílios com erro
Total	**5 304 711**	**1 494 397**	**28,17**
1 pessoa	497 673	63 571	12,77
2 pessoas	873 061	157 554	18,05
3 pessoas	1 134 064	264 849	23,35
4 pessoas	1 210 216	353 623	29,22
5 pessoas	759 139	266 754	35,14
6 pessoas	388 930	164 640	42,33
7 pessoas	192 825	91 115	47,25
8 pessoas	108 764	56 367	51,83
9 ou mais pessoas	140 039	75 905	54,20

Fonte: IBGE, Censo Demográfico 2000.

Nota: Ocorreram 19 casos de registros pertencentes aos estratos de 1 até 8 pessoas que foram detectados como possuidores de algum erro ou inconsistência, mas para os quais não foi possível encontrar doador pelo sistema NIM. Esses 19 casos foram a valiados e imputados pelo sistema IMPS com os 75 905 casos do estrato de 9 ou mais pessoas

Neste caso, é razoável esperar que a percentagem de erros seja proporcional ao tamanho do estrato (quanto maior o estrato, maior será a percentagem de erros). A importância desta informação e, em especial, a relativa ao estrato de 9 ou mais pessoas, é que ela indica a quantidade de domicílios com erro que se pode esperar para serem corrigidos manualmente, caso esta estratégia seja, de novo, usada Ou seja, neste último estrato, um pouco mais da metade dos questionários tem algum erro de estrutura e, na hora de definir os lotes de produção, deve-se levar em conta que somente 50% dos questionários poderiam ser considerados como doadores em potencial.

Gráfico 12.15 - Totais de erros, segundo os estratos



A análise de erros (imputações) por domicílio é apresentada na Tabela 12.22, e mostra que 60% dos domicílios teve somente uma variável imputada (esta percentagem teria sido muito maior se houvesse uso de uma estratégia de imputação automática do número da família (ver mais adiante). Por curiosidade, pode-se ver que, ao final da lista, aparece um domicílio com 43 variáveis imputadas.

Tabela 12.22 - Número de domicílios, por total de erros por domicílio

(continua)

Número de erros por domicílio	Freqüência absoluta	Freqüência relativa (%)	Freqüência relativa acumulada (%)
Total	**1 418 473**	**100,00**	**100,00**
1	848 007	59,78	59,78
2	275 299	19,41	79,19
3	124 473	8,78	87,97
4	69 938	4,93	92,90
5	37 404	2,64	95,53
6	23 212	1,64	97,17
7	13 621	0,96	98,13
8	8 912	0,63	98,76
9	5 576	0,39	99,15
10	3 815	0,27	99,42
11	2 543	0,18	99,60
12	1 785	0,13	99,73
13	1 205	0,08	99,81
14	796	0,06	99,87
15	588	0,04	99,91

Tabela 12.22 - Número de domicílios, por total de erros por domicílio

(conclusão)

Número de erros por domicílio	Freqüência absoluta	Freqüência relativa (%)	Freqüência relativa acumulada (%)
16	432	0,03	99,94
17	251	0,02	99,96
18	154	0,01	99,97
19	117	0,01	99,98
20	94	0,01	99,98
21	63	0,00	99,99
22	49	0,00	99,99
23	32	0,00	99,99
24	25	0,00	99,99
25	26	0,00	100,00
26	9	0,00	100,00
27	13	0,00	100,00
28	9	0,00	100,00
29	4	0,00	100,00
30	4	0,00	100,00
31	5	0,00	100,00
32	3	0,00	100,00
33	1	0,00	100,00
34	1	0,00	100,00
35	5	0,00	100,00
37	1	0,00	100,00
43	1	0,00	100,00

Fonte: IBGE, Censo Demográfico 2000.

Evidentemente que o número de variáveis imputadas deve ser proporcional ao estrato, ou melhor, ao número de pessoas por domicílio. A Tabela 12.23 apresenta a razão de erros por pessoa, e pode-se ver que quase 58% dos domicílios têm menos de ½ erro por pessoa, e 73% têm até ½ erro por pessoa.

Tabela 12.23 - Número de domicílios, segundo a razão de erros por pessoa

Razão de erros por pessoa	Número de domicílios com erros	Percentual de domicílios com erros
Total	**1 418 473**	**100,00**
0. Menos de 0,5 erros por pessoa	819 726	57,79
1. 0,5 erros por pessoa	222 321	15,67
2. Mais de 0,5 até 1 erro por pessoa	285 141	20,10
3. Mais de 1 erro até 2 erros por pessoa	77 721	5,48
4. Mais de 2 erros por pessoa	13 564	0,96

Fonte: IBGE, Censo Demográfico 2000.

A Tabela 12.24 mostra a percentagem de imputação por variável. Vê-se que quase 1/3 (33%) do total de imputações corresponde à variável V0404 "número da família". A outra grande freqüência concentra-se na V0403 "relação com o responsável pela família", com cerca de 24%. Isto demonstra que não foi fácil responder a estas duas perguntas e que, provavelmente, seja necessário rever estes conceitos e/ou estas perguntas, nas próximas pesquisas.

Tabela 12.24 - Freqüência de imputação, segundo a variável

Variável	Número de imputações	Percentual de imputações
Total	**2 742 154**	**100,00**
V0401	247 174	9,01
V0402	227 268	8,29
V0403	647 312	23,61
V0404	907 712	33,10
V0436	248 873	9,08
V0437	218 168	7,96
V0438	123 565	4,51
V4007	5 367	0,20
V4667	116 715	4,26

Fonte: IBGE, Censo Demográfico 2000.

No caso da variável V0404, a Tabela 12.25 mostra a distribuição de freqüências dos valores errados (antes da imputação). A categoria 10 significa "valor em branco", e responde por quase 68% dos valores errados. Em outras palavras, 68% das pessoas que tiveram a variável V0404 imputada, o foram porque simplesmente estava em branco (não informada).

Tabela 12.25 - Freqüência dos valores errados da variável V0404, segundo seus valores antes da imputação

Valor da variável V0404 antes da imputação	Freqüência	Freqüência relativa (%)
Total	**907 712**	**100,00**
0	3 953	0,44
1	35 556	3,92
2	208 248	22,94
3	30 706	3,38
4	7 047	0,78
5	2 760	0,30
6	1 539	0,17
7	975	0,11
8	336	0,04
9	143	0,02
Branco	616 449	67,91

Fonte: IBGE, Censo Demográfico 2000.

Mais curioso, ainda, é que na Tabela 12.26 – matriz de imputação da variável V0404 – pode-se ver que a maioria dos valores em branco (606.128 de 616.449, ou 98%) foram imputados para o valor 1 (família 1). Isto era espera-

do, visto que 93% dos domicílios têm somente uma família. O importante é que esta variável poderia ser objeto de uma pré-imputação, transformando, por exemplo, todos os brancos no valor 1. Isto facilitaria os trabalhos de imputação, porque, possivelmente, estes domicílios seriam considerados corretos e aumentariam a massa de domicílios doadores em cada estrato, aumentando, assim, a chance de termos menores distâncias e uma qualidade maior na imputação.

Tabela 12.26 - Matriz de imputação da variável V0404: valores antes e depois da imputação

Valor da variável V0404 antes da imputação	Valor da variável V0404 depois da imputação					
	Total	0	1	2	3	4
Total	**3 953**	-	**3 918**	**35**	-	-
0	35 556	5 000	-	29 473	1 077	6
1	208 248	20	206 437	-	1 777	14
2	30 706	4	20 177	10 510	-	15
3	7 047	5	6 537	401	104	-
4	2 760	-	2 672	84	3	1
5	1 539	2	1 468	67	2	-
6	975	-	924	50	1	-
7	336	-	314	22	-	-
8	143	-	133	10	-	-
9	616 449	1 056	606 128	8 994	269	2
Branco	907 712	6 087	848 708	49 646	3 233	38

Fonte: IBGE, Censo Demográfico 2000.

Por outro lado, no caso da variável V0403 "relação com o responsável pela família", a Tabela 12.27 apresenta a distribuição de freqüências dos valores errados. A categoria 0 significa "valor em branco" e responde por 17% desses valores. O interessante aqui é a percentagem de casos imputados quando o entrevistado se declarou responsável pela família (categoria 1, com 48%). Isto significa que quase metade das imputações nesta variável foram causadas porque alguém respondeu que a pessoa era responsável pela família. Pelos conceitos do programa de imputação isto foi considerado errado.

Tabela 12.27 - Freqüência dos valores errados da variável V0403, segundo seus valores antes da imputação

Valor da variável V0403 antes da imputação	Freqüência	Freqüência relativa (%)
Total	**647 312**	**100,00**
1	314 325	48,56
2	81 259	12,55
3	78 823	12,18
4	9 615	1,49
5	4 916	0,76
6	6 785	1,05
7	29 688	4,59
8	5 257	0,81
9	870	0,13
10	1 698	0,26
11	1 432	0,22
12	486	0,08
Branco	112 158	17,33

Fonte: IBGE, Censo Demográfico 2000.

Ao verificar-se a Tabela 12.28, que mostra a matriz de imputação desta variável, pode-se ver que a grande maioria dos "supostos responsáveis" foi imputada para a categoria 2 (cônjuge), ou seja, 270.183 de 314.325. Possivelmente, isto pode significar que estas pessoas se considerariam dividindo a responsabilidade da família, o que seria uma razão a mais para uma revisão profunda destes conceitos.

Tabela 12.28 - Matriz de imputação da variável V0403: valores antes e depois da imputação

Valor da variável V0403 antes da imputação	Valor da variável V0403 depois da imputação												
	Total	1	2	3	4	5	6	7	8	9	10	11	12
Total	**647 312**	**107 146**	**334 709**	**146 069**	**7 360**	**21 359**	**7 091**	**14 650**	**1 861**	**538**	**825**	**21**	**5 683**
1	314 325	-	270 183	26 852	3 395	748	2 482	4 596	711	264	327	1	4 766
2	81 259	21 569	-	52 703	1 458	431	843	3 992	154	29	77	2	1
3	78 823	42 763	15 935	-	680	15 754	1 269	2 143	177	19	59	15	9
4	9 615	5 476	3 427	640	-	30	16	17	1	-	1	-	7
5	4 916	1 374	544	2 545	8	-	86	334	17	1	6	-	1
6	6 785	3 063	1 626	1 420	22	120	-	436	62	13	19	-	4
7	29 688	12 466	13 095	2 875	265	346	394	-	176	17	39	1	14
8	5 257	1 601	1 569	1 106	27	73	118	605	-	31	60	-	67
9	870	276	180	149	4	11	55	66	33	-	1	-	95
10	1 698	826	278	396	2	20	27	80	53	4	-	-	12
11	1 432	47	55	1 096	1	65	53	73	22	2	18	-	-
12	486	419	12	27	1	1	18	5	3	-	-	-	-
Branco	112 158	17 266	27 805	56 260	1 497	3 760	1 730	2 303	452	158	218	2	707

Fonte: IBGE, Censo Demográfico 2000.

Seguindo com a análise por variáveis, a variável V0402 “relação com o responsável pelo domicílio” não apresenta maiores peculiaridades, como mostra Tabela 12.29, a não ser a maior freqüência de imputados por falta de informação (representada pela categoria 0) em 25%.

Tabela 12.29 - Freqüência dos valores errados da variável V0402, segundo seus valores antes da imputação

Valor da variável V0402 antes da imputação	Freqüência	Freqüência relativa (%)
Total	**227 268**	**100,00**
1	15 526	6,83
2	32 172	14,16
3	33 702	14,83
4	2 569	1,13
5	34 248	15,07
6	8 708	3,83
7	23 149	10,19
8	11 976	5,27
9	2 244	0,99
10	2 308	1,02
11	2 249	0,99
12	527	0,23
Branco	57 890	25,47

Fonte: IBGE, Censo Demográfico 2000.

A análise da variável V0401 "sexo", cuja distribuição de freqüência está na Tabela 12.30, mostra que 68% das imputações são devidas à falta de informação (categoria 0): quase 170 mil respostas em branco na variável "sexo" que, em princípio, é de fácil preenchimento. Este número não representa muito no total de pessoas da amostra (um pouco mais de 20 milhões), mas é sintomático de que é preciso mais atenção para as próximas pesquisas.

Tabela 12.30 - Freqüência dos valores errados da variável V0401, segundo seus valores antes da imputação

Valor da variável V0401 antes da imputação	Freqüência	Freqüência relativa (%)
Total	**247 174**	**100,00**
1	35 151	14,22
2	42 627	17,25
Branco	169 396	68,53

Fonte: IBGE, Censo Demográfico 2000.

Pela Tabela 12.23, vê-se que a variável V0436 "vive em companhia de cônjuge ou companheiro" apresentou cerca de 9% das imputações realizadas. A Tabela 12.31 apresenta a freqüência absoluta e relativa dos valores possíveis da variável antes da imputação e a Tabela 12.32 apresenta, em valores percentuais, a matriz com a distribuição dos valores possíveis da variável antes e depois do processo de imputação.

Tabela 12.31 - Freqüência dos valores errados da variável V0436, segundo seus valores antes da imputação

Valor da variável V0436 antes da imputação	Freqüência	Freqüência relativa (%)
Total	**248 873**	**100,00**
Sim	151 391	60,83
Não, mas viveu	30 645	12,31
Nunca viveu	9 919	3,99
Branco	56 918	22,87

Fonte: IBGE, Censo Demográfico 2000.

Tabela 12.32 - Matriz de imputação da variável V0436: valores percentuais antes e depois da imputação

Valor da variável V0436 antes da imputação	Valor das variáveis V0436 depois da imputação				
	Total	Sim	Não, mas viveu	Nunca viveu	Branco
Total	**100,00**	**16,57**	**44,64**	**38,68**	**0,10**
Sim	60,83	-	41,00	19,75	0,08
Não, mas viveu	12,31	9,01	-	3,30	0,01
Nunca viveu	3,99	2,88	1,09	-	0,02
Branco	22,87	4,69	2,55	15,63	-

Fonte: IBGE, Censo Demográfico 2000.

Na Tabela 12.31, destaca-se o fato de 60% das imputações terem acontecido em registros de pessoas que responderam o código 1 – Sim, ou seja, informaram que viviam em companhia de cônjuge ou companheiro, mas que as informações das demais pessoas do domicílio não confirmavam essa informação. O mais importante ainda, apresentado na Tabela 12.31, é o fato de que do total das imputações realizadas nessa variável, 40% referem-se à troca do código 1 – Sim para o código 2 – Não, mas já viveu. Isso pode indicar algum viés na coleta da informação e carece de estudos para uma melhor compreensão da ocorrência.

O processamento através do sistema IMPS

No conjunto de críticas que seriam executadas por esse sistema, verificou-se que algumas eram determinísticas e que poderiam ser processadas através de programa. O documento interno *Procedimentos a serem implementados para a execução da crítica (NIM e IMPS), questionário da amostra/censo 2000* (2001), especifica as situações que deveriam ser corrigidas antes de se executar o IMPS.

Os procedimentos de correção determinística ocorreram em situações muito específicas, não esgotando, portanto, todos os acertos nessas variáveis, abrangendo as seguintes variáveis: sexo; relação com a pessoa responsável pelo domicílio e pela família; número da família; e a variável auxiliar faixa de idade, apenas na categoria inválido.

Para a correção determinística relativa à variável "sexo", nos casos de omissão do registro para as pessoas de dez anos e mais de idade, foram consideradas as especificações da variável auxiliar "marca da fecundidade". Já para as pessoas menores de dez anos de idade, foi desenvolvido um programa de imputação probabilística. Informações detalhadas sobre esse sistema podem ser conhecidas através do documento *Manual de crítica: sistema IMPS* (2002).

Ao submeter o lote de trabalho ao IMPS, o operador observava na tela do microcomputador as mensagens de erro, identificadas de acordo com o conjunto de críticas elaboradas para o tratamento das inconsistências entre registros. As correções só eram realizadas após a análise criteriosa de todas as variáveis envolvidas nas críticas entre registros, disponibilizadas para o operador. Alguns desses critérios, por razões de confiabilidade na variável, orientavam o operador a só alterar as variáveis faixa de idade e a marca da fecundidade em casos estritamente necessários. Por razões operacionais, na execução do sistema IMPS os domicílios com mais de 38 moradores não puderam ser criticados em tela do microcomputador. Esses domicílios, um de Pernambuco com 42 moradores e o outro de Sergipe com 43 moradores, foram criticados e corrigidos separadamente.

A equipe que esteve envolvida no trabalho de correção através do sistema IMPS era composta por treze técnicos. Em média eram corrigidos 600 questionários por dia. Os 67 lotes que foram criados e submetidos ao IMPS totalizavam 140.029 questionários com erros. Os lotes foram submetidos ao sistema de crítica a partir do dia 10/04/2002, tendo-se concluído toda a correção no dia 22/05/2002.

Após analisar cada mensagem de erro e solucionar o problema, questionário a questionário, o operador acionava um comando que submetia todo lote novamente ao sistema de crítica para certificar-se de que não apresentava mais erros.

Tendo em vista a crítica e imputação da variável "sexo", os totais desta variável – homens e mulheres – de cada domicílio, foram recalculados.

12.5.2.2 Crítica intra-registros

Antes da descrição dos diversos aplicativos utilizados para detecção e depuração das inconsistências dos dados do Questionário da Amostra, é necessário ter conhecimento de algumas questões.

Em primeiro lugar, devido a complexidade dos temas investigados no Questionário da Amostra e, consequentemente, do número de variáveis que o compõe, só foi possível implementar o processo de imputação utilizando o sistema DIA, a partir da criação de um conjunto bastante significativo de variáveis auxiliares, antes de se iniciar a crítica em cada aplicativo.

Considerando-se que essas variáveis auxiliares foram criadas, inicialmente, com base em arquivos ainda não corrigidos, houve a necessidade, assim que se concluía a imputação de um aplicativo, que os totais fossem gerados novamente, a partir dos arquivos depurados.

Por outro lado, a análise do resultado da imputação, realizada durante a execução dos testes dos aplicativos com os dados do censo, para o Espírito Santo e Rondônia, apontou algumas situações de inconsistência. Essas situações não decorreram de equívocos na execução do DIA. Ou seja, na sua essência, o processo de imputação funcionou de acordo com os pressupostos do sistema, gerando registros que obedeciam ao estabelecido pelas regras de crítica. O que se detectou foi uma perda na coerência entre algumas poucas variáveis. Por esse motivo, decidiu-se implementar alguns procedimentos prévios à execução dos aplicativos do DIA – pré-DIA, com vistas a preparar essas variáveis para serem imputadas.

Assim, no bloco de fecundidade, foram implementados procedimentos pré-DIA em alguns registros envolvendo as variáveis: filhos tidos nascidos vivos, filhos tidos nascidos vivos que estavam vivos e filhos tidos nascidos mortos. No bloco de instrução, foram previamente tratados alguns registros relativos às variáveis “conclusão do curso no qual estudou” e “código do curso mais elevado concluído”, bem como, no bloco de mão-de-obra, alguns dos que envolviam a variável “posição na ocupação”. A variável auxiliar “idade calculada do entrevistado”, obtida através de um algoritmo próprio, também passou por esse tipo de ajuste prévio. Todos os procedimentos pré-DIA podem ser consultados no documento Definições necessárias à implementação da apuração centralizada dos dados referentes ao questionário da amostra (CD 1.02) no censo demográfico de 2000 (2003).

A definição do conjunto de aplicativos para a imputação dos dados da amostra obedeceu, em primeiro lugar, à conveniência da composição em um só aplicativo das críticas relativas a blocos distintos do questionário, buscando-se estabelecer um processo de imputação mais criterioso. Em função disso, foi implementada a imputação conjunta de parte dos blocos de instrução e fecundidade.

O segundo aspecto levado em conta para o estabelecimento dos aplicativos, foi garantir ao DIA a possibilidade de geração do conjunto completo de críticas, tendo em conta o elevado número de variáveis a serem depuradas no Questionário da Amostra. Desse modo, os aplicativos foram definidos, levando-se em conta a quantidade de variáveis investigadas em cada parte do questionário, mas também o número de regras de crítica a serem implementadas. Isso determinou a partição dos aplicativos de instrução e fecundidade, de mão-de-obra, de migração e de fecundidade.

Além disso, também foi importante a definição da ordem de execução dos aplicativos. A experiência do passado e a análise dos resultados da imputação, durante a fase de testes, mostrou a importância de determinadas variáveis serem imputadas somente após a correção prévia de outras. Dentro dessa lógica, certas variáveis já corrigidas tornar-se-iam fixas durante a execução dos aplicativos seguintes.

Já no caso dos aplicativos de fecundidade, além de todas as razões apontadas, as divisões I, II e III ocorreram, tendo em vista a necessidade de se estabelecer uma ordenação na seqüência em que as variáveis deveriam ser imputadas, a partir do grau de confiança associado às variáveis envolvidas.

As variáveis do bloco de fecundidade foram imputadas, em parte, nos aplicativos Instrução e Fecundidade da Pessoa Responsável pela Família ou Individual em Domicílio Coletivo e Instrução e Fecundidade das Demais Pessoas da Família e, em parte, nos aplicativos Fecundidade I, II e III.

A ordem de execução dos aplicativos no sistema DIA foi a seguinte: Instrução e Fecundidade da Pessoa Responsável pela Família ou Individual em Domicílio Coletivo; Instrução e Fecundidade das Demais Pessoas da Família; Mão-de-Obra I e II; Migração I e II; Deficiência; e Fecundidade I, II e III.

Após os lotes terem sido imputados através desse aplicativos, alguns procedimentos de conclusão foram necessários – procedimentos pós-DIA. Como exemplo, temos o decorrente da estratégia utilizada durante a execução do sistema, ou seja, a necessidade de submeter os lotes a um tradutor para os códigos de ocupação e atividade, recompondo o banco original codificado a cinco dígitos, visto que o sistema de crítica foi executado com os códigos a três dígitos. Outros procedimentos pós-Dia, decorrentes do processo analítico da variável Rendimento, também foram implementados.

Os procedimentos desenvolvidos no pós-DIA, podem ser conhecidos, consultando-se as Referências, no final do capítulo.

a) Aplicativo Domicílio

As variáveis do questionário imputadas por este aplicativo foram:

- V0202 – Tipo;

- V0203 – Total de cômodos;

- V0204 – Cômodos servindo de dormitório;

- V0205 – Condição de ocupação do domicílio;

- V0206 – Condição de ocupação do terreno do domicílio;

- V0207 – Forma de abastecimento de água;

- V0208 – Canalização da água;

- V0209 – Número de banheiros;

- V0210 – Existência de sanitário;

- V0211 – Tipo de escoadouro;

- V0212 – Destino do lixo;

- V0213 – Iluminação elétrica;

- V0214 – Existência de rádio;

- V0215 – Existência de geladeira ou freezer;

- V0216 – Existência de videocassete;

- V0217 – Existência de máquina de lavar roupa;

- V0218 – Existência de forno de microondas;

- V0219 – Existência de linha telefônica instalada;

- V0220 – Existência de microcomputador;

- V0221 – Quantidade de televisores;

- V0222 – Quantidade de automóveis para uso particular; e

- V0223 – Quantidade de aparelhos de ar condicionado.

Durante a execução do aplicativo, as variáveis V0201 "espécie do domicílio" e V7100 "total de pessoas no domicílio foram declaradas fixas, servindo de referência para a correção das demais variáveis.

Todas as variáveis foram imputadas probabilisticamente, de acordo com as suas respectivas distribuições marginais dadas pelas freqüências dos registros não suspeitos e o método proporcional.

A variável V7100 foi utilizada, especificamente, para corrigir as incompatibilidades envolvendo a variável V0204. As variáveis V0205 e V0209 foram tratadas com pesos, respectivamente, 2 e 1, enquanto às demais foi atribuído o peso médio igual a 5.

b) Aplicativo Instrução e Fecundidade da Pessoa Responsável pela Família ou Individual em Domicílio Coletivo

Neste aplicativo foram corrigidas as inconsistências relativas às variáveis do CD 1.02, de todo o bloco de instrução e algumas relativas ao bloco de fecundidade, somente para a "pessoa responsável pela família" ou classificada como "individual em domicílio coletivo".

As variáveis criadas e imputadas, foram:

- V4620 – Total de filhos tidos nascidos vivos;

-V0463 – Total de filhos tidos nascidos vivos e que estavam vivos;

- V4640 – Indicadora do preenchimento da V0464;

- V4660 – Indicadora do preenchimento da V0466;

- V4670 – Total de filhos nascidos mortos;

- V4075 – Idade calculada do entrevistado; e

- V4654 – Idade calculada do último filho nascido vivo.

As seguintes variáveis do questionário foram imputadas:

- V0428 – Sabe ler e escrever;

- V0429 – Freqüência à escola ou creche;

-- V0430 – Curso que freqüenta;

- V0431 – Série que freqüenta;

- V0432 – Curso mais elevado que freqüentou no qual concluiu pelo menos uma série;

- V0433 – Última série concluída com aprovação;

- V0434 – Conclusão do Curso no Qual Estudou; e

- V4353 – Código do curso mais elevado concluído;

As variáveis auxiliares, criadas e consideradas fixas, foram as seguintes:

- V4060 – Grupo de idade do cônjuge;

- V4301 – Grupo de anos de estudo;

- V4403 – Indicadora do estado conjugal da mulher;

- V4567 – Grupo de idade qüinqüenal do entrevistado; e

- V4453 – Código, a três dígitos, da variável ocupação.

Foram as seguintes, as variáveis tratadas em etapas anteriores, portanto fixas:

- V4007 – Faixa de idade do entrevistado;

- V0401 – Sexo;

- V0402 – Relação com a pessoa responsável pelo domicílio; e

- V0403 – Relação com a pessoa responsável pela família.

Para o bloco de fecundidade, foram corrigidas as inconsistências relativas ao "total de filhos tidos nascidos vivos", ao "total de filhos tidos nascidos vivos e que estavam vivos" e ao "total de filhos nascidos mortos", sem a intenção de se corrigir as suas parcelas – homens e mulheres – que foram tratadas em aplicativo posterior.

O procedimento de imputação implementado para esse aplicativo, centrou-se na utilização de distribuição conjunta dos registros não suspeitos e no método proporcional, para as seguintes variáveis:

- V0428 conjunta com a variável V4075;

- V4075 conjunta com as variáveis V0403 e V4060;

- V4620 conjunta com as variáveis V0463 e V4567;

- V0463 conjunta com as variáveis V4301 e V4567;

- V4654 conjunta com as variáveis V4620 e V4567; e

- V4670 conjunta com as variáveis V4567 e V4301.

As variáveis V0430 e V0431 foram imputadas no DIA, de acordo com a situação, através de método determinístico ou probabilístico. A imputação determinística passou a ser uma estratégia em virtude da ocorrência de eventuais inconsistências entre os registros da série e do grau, envolvendo níveis de ensino diferentes: primário, ginásio e colegial, clássico, científico etc. – quando o correto seria primeiro grau ou segundo grau. Nos demais casos, em que a modalidade foi a imputação probabilística, utilizou-se a distribuição marginal dos registros não suspeitos e o método proporcional.

Para todas as demais variáveis tratadas nesse aplicativo, a imputação aconteceu de acordo com a distribuição marginal dos registros não suspeitos, através do método proporcional. A ponderação para a variável V4075, devido ao maior grau de confiabilidade em relação às demais variáveis, foi 1 e, para todas as outras, utilizou-se o peso médio 5.

c) Aplicativo Instrução e Fecundidade das Demais Pessoas da Família

Neste aplicativo, foram corrigidas as inconsistências relativas às variáveis do CD 1.02 de todo o bloco de instrução e algumas do bloco de fecundidade, das demais pessoas da família, exceto a "pessoa responsável" ou o "individual em domicílio coletivo".

Os procedimentos utilizados para a imputação foram, quase todos, os mesmos utilizados e descritos no aplicativo anterior. Desse modo, as variáveis tratadas em etapas anteriores e as imputadas, do questionário ou criadas, foram exatamente as mesmas.

Da lista de variáveis criadas e consideradas fixas do aplicativo anterior, foi excluída a V4060 e introduzida a V4061"grupo de idade da pessoa responsável pela família" ou "individual em domicílio coletivo". Assim, essa nova variável passou a ser considerada, através da distribuição conjunta, para fazer a imputação da variável V4075 "idade calculada do entrevistado".

d) Aplicativos de Mão-de-Obra

Nestes aplicativos, foram tratadas as inconsistências relativas ao bloco de mão-de-obra do CD 1.02. Houve a necessidade de implementar dois aplicativos, Mão-de-Obra I e Mão-de-Obra II.

Aplicativo Mão-de-Obra I

Neste aplicativo, as variáveis criadas e imputadas foram:

- V4453 – Código, a três dígitos, da variável ocupação;

- V4510 – Indicadora de rendimento no trabalho principal;

- V4520 – Indicadora de rendimento nos demais trabalhos;

- V4570 – Indicadora de rendimento de aposentadoria ou pensão;

- V4580 – Indicadora de rendimento de aluguel;

- V4590 – Indicadora de rendimento de pensão alimentícia, mesada ou doação recebida de não-morador;

- V4600 – Indicadora de rendimento de renda mínima, bolsa-escola, seguro-desemprego etc.; e

- V4610 – Indicadora de outros rendimentos.

As variáveis imputadas do questionário foram:

- V0439 – Trabalho remunerado na semana de referência;

-V0440 – Trabalho remunerado do qual estava temporariamente afastado na semana de referência;

- V0441 – Ajuda sem remuneração na semana de referência no trabalho exercido por pessoa moradora do domicílio ou trabalho sem remuneração como aprendiz ou estagiário;

- V0442 – Ajuda sem remuneração na semana de referência no trabalho exercido por pessoa moradora do domicílio empregada em atividade de cultivo, extração vegetal etc.;

- V0443 – Trabalho na semana de referência em atividade de cultivo, extração vegetal0 etc. destinados à alimentação de pessoas moradoras do domicílio;

- V0444 – Número de trabalhos na semana de referência;

- V0447 – Posição na ocupação;

- V0448 – Empregado pelo regime jurídico dos funcionários públicos ou como militar;

- V0449 – Número de empregados;

- V0450 – Contribuinte de instituto de previdência oficial;

- V0453 – Número de horas trabalhadas por semana no trabalho principal;

- V0454 – Número de horas trabalhadas por semana nos demais trabalhos;

- V0455 – Providência para conseguir trabalho; e

- V0456 – Aposentado de instituto de previdência oficial.

Foram as seguintes, as variáveis auxiliares criadas e consideradas fixas:

- V4301 – Grupo de anos de estudo;

- V4302 – Grupo de anos de estudo, por sexo;

- V4471 – Grupo de Posição na Ocupação;

- V4568 – Indicadora de grupos de idade a partir de 10 anos.

Foram consideradas fixas, as seguintes variáveis tratadas em etapas anteriores:

- V4075 – Idade calculada do entrevistado;

- V0428 – Sabe ler e escrever;

- V0429 – Freqüência à escola ou creche;

- V0430 – Curso que freqüenta;

- V4353 – Código do Curso Mais Elevado Concluído; e

- V0402 – Relação com a pessoa responsável pelo domicílio.

À exceção da variável V0447, que teve peso igual a 1, todas as demais do Aplicativo Mão-de-Obra I foram imputadas com peso médio 5.

Para a imputação das variáveis, V0439 a V0443 e V0455 foram utilizados os registros não suspeitos e o método proporcional da distribuição conjunta com a variável V4568. Já a estratégia de imputação das variáveis V4580, V4590, V4600 e V4610 considerou a distribuição conjunta com a variável V4301. Por fim, para a variável V4453 foi levada em conta a distribuição conjunta com as variáveis V4302 e V4471.

As demais variáveis do aplicativo foram imputadas de acordo com a distribuição marginal dos registros não suspeitos, através do método proporcional.

Aplicativo Mão-de-Obra II

Neste aplicativo, foram corrigidas as inconsistências na variável V4463 "Código a Três Dígitos da Variável Atividade". Para isso, as variáveis do Aplicativo Mão-de-Obra I, que serviram de referência, permanecendo fixas, foram a V4453 e V0448. Como ponderação, atribuiu-se o peso médio 5 e a variável V4463 foi imputada de acordo com a distribuição marginal dos registros não suspeitos, através do método proporcional.

e) Aplicativos de Migração

Nestes aplicativos, foram corrigidas as inconsistências do bloco de migração do CD 1.02 para todas as pessoas do domicílio.

Todas as variáveis, considerando-se os dois aplicativos implementados, foram imputadas de acordo com a distribuição marginal dos registros não suspeitos, através do método proporcional. Com exceção da variável V4254, que teve ponderação 3, todas as demais variáveis tiveram peso médio 5.

Aplicativo Migração I

Para este aplicativo, as variáveis criadas e imputadas foram as seguintes:

- V4201 – Indicadora de preenchimento da V0420.

- V4254 – Indicadora da relação de igualdade entre a V4251 e a V0102;

- V4264 – Recodificação da V4261;. e

- V4270 – Recodificação da V4276.

As variáveis imputadas, originais do Questionário da Amostra, foram:

- V0415 – Sempre morou neste município;

- V0416 – Tempo de moradia sem interrupção neste município;

- V0417 – Nasceu neste município;

- V0418 – Nasceu nesta Unidade da Federação;

- V0419 – Nacionalidade;

- V4210 – Código da Unidade da Federação ou país estrangeiro de nascimento;

- V0422 – Tempo de moradia sem interrupção na Unidade da Federação;

- V4230 – Código da Unidade da Federação ou país estrangeiro de residência anterior; e

- V0424 – Local de Residência em 31 de Julho de 1995.

As variáveis auxiliares criadas e consideradas fixas foram:

- V1023 – Indicadora de UF/município totalmente urbano, onde foi realizado o censo.

- V1030 – Indicadora de município de Brasília;

- V4454 – Indicadora de preenchimento da ocupação;

- V4231 – Indicadora da relação entre a V4230 e V0102; e

- V4252 – Indicadora de município totalmente urbano.

As variáveis tratadas em etapas anteriores e, portanto, consideradas fixas, foram:

- V4075 – Idade calculada do entrevistado; e

- V0429 – Freqüência à escola ou creche.

Aplicativo Migração II

No aplicativo Migração II, foram corrigidas todas as inconsistências do bloco de migração de todas as pessoas do domicílio.

Houve a criação da variável V4251 "Indicadora da UF/país estrangeiro", e as variáveis do questionário, objeto de imputação, foram as seguintes:

- V0420 – Ano em que fixou residência no Brasil; e

- V4260 – Código da UF ou país estrangeiro de residência em 31/7/95.

As variáveis tratadas em etapas anteriores e, portanto, fixas, foram:

- V4075 – Idade calculada do entrevistado;

- V0419 – Nacionalidade;

- V4201 – Indicadora de preenchimento da V0420;

- V0422 – Tempo de moradia sem interrupção na Unidade da Federação;

- V0102 – Código da UF onde foi realizado o censo;

- V4254 – Indicadora da relação de igualdade entre a V4251 e a V0102 ; e

- V4264 – Recodificação da V4261.

f) Aplicativo Deficiência

Neste aplicativo, foram corrigidas, para todos os moradores do domicílio, as inconsistências entre as variáveis do bloco de deficiência e as incompatibilidades entre essas variáveis e a variável "ocupação". As omissões de registro para as variáveis "cor ou raça" e "religião ou culto" foram corrigidas no DIA através de imputação determinística, atribuindo-lhes o código de informação ignorada.

As variáveis imputadas, pertencentes ao questionário, foram:

- V0408 – Cor ou raça;

- V4090 – Código da religião ou culto;

- V0410 – Deficiência mental permanente;

- V0411 – Autoavaliação da capacidade de enxergar;

- V0412 – Autoavaliação da capacidade de ouvir;

- V0413 – Autoavaliação da capacidade de caminhar / subir escadas; e

- V0414 – Existência de deficiências

As variáveis tratadas em etapas anteriores e, portanto, fixas, foram as seguintes:

- V0402 – Relação com a pessoa responsável pelo domicílio;

- V0403 – Relação com a pessoa responsável pela família;

- V4007 – Faixa de idade do entrevistado; e

- V4453 – Código, a três dígitos, da variável ocupação.

As variáveis tiveram o peso médio 5 e utilizou-se para a imputação a distribuição marginal dos registros não suspeitos através do método proporcional.

g) Aplicativos de Fecundidade

Através dos Aplicativos Fecundidade I, II e III, concluiu-se a correção de todo o bloco de fecundidade, iniciada pelos dois Aplicativos de Instrução e Fecundidade, descritos anteriormente. O peso médio 5 foi atribuído às variáveis nesses três aplicativos. Parte da imputação do aplicativo Fecundidade III - para mulheres com mais de 19 filhos - foi feita através de imputação manual, pois o DIA não conseguiu criar o conjunto completo de críticas

Aplicativo Fecundidade I

Neste aplicativo, foram corrigidas as inconsistências em relação às características da fecundidade das mulheres cuja variável V4620 "total de filhos tidos nascidos vivos" estava em branco ou preenchida com a informação 0 (zero) até 8. Foram também imputadas, as parcelas, homens e mulheres, da variável V4670 "total de filhos tidos nascidos mortos" para todas as mulheres.

Foram as seguintes as variáveis imputadas, investigadas no questionário:

- V4621 – Filhos tidos nascidos vivos, homens;

- V4622 – Filhos tidos nascidos vivos, mulheres;

- V4631 – Filhos tidos nascidos vivos que estavam vivos, homens;

- V4632 – Filhos tidos nascidos vivos que estavam vivos, mulheres;

- V0464 – Sexo do último filho nascido vivo;

- V4671 – Filhos tidos nascidos mortos, homens; e

- V4672 – Filhos tidos nascidos mortos, mulheres.

As variáveis tratadas em etapas anteriores, portanto fixas, foram:

- V4620 – Total de filhos tidos nascidos vivos;

- V0463 – Total de filhos tidos nascidos vivos que estavam vivos; e

- V4670 – Total de filhos tidos nascidos mortos.

Como estratégia, algumas variáveis foram imputadas de acordo com a distribuição conjunta, através dos registros não suspeitos, e o método proporcional, a saber:

- as variáveis V4671 e V4672, através da variável V4670;

- as variáveis V4621 e V4622, através da variável V4620; e

- as variáveis V4631 e V4632, através da variável V0463.

As demais variáveis desse aplicativo foram imputadas de acordo com a distribuição marginal dos registros não suspeitos e o método proporcional.

Aplicativo Fecundidade II

No Aplicativo Fecundidade II, foram corrigidas as inconsistências em relação às variáveis V4631 e V4632, das mulheres cuja variável V4620 assumiu um valor de 9 a 31.

As variáveis imputadas, presentes no questionário, foram:

- V4631 – Filhos tidos nascidos vivos que estavam vivos, homens; e

- V4632 – Filhos tidos nascidos vivos que estavam vivos, mulheres.

As variáveis tratadas em etapas anteriores, portanto fixas, foram:

- V4620 – Total de filhos tidos nascidos vivos; e

- V0463 – Total de filhos tidos nascidos vivos que estavam vivos.

Utilizou-se a distribuição conjunta dos registros não suspeitos e o método proporcional para imputar as variáveis V4631 e V4632, por meio da V0463

Aplicativo Fecundidade III

Neste aplicativo, foram corrigidas as inconsistências em relação às características da fecundidade relativas às variáveis V4621, V4622 e V0464, das mulheres cuja variável V4620 estava preenchida com valores de 9 a 19.

As variáveis V4621 e V4622 foram imputadas de acordo com a variável V4620, através da distribuição conjunta dos registros não suspeitos e o método proporcional. As demais variáveis deste aplicativo foram imputadas de acordo com a distribuição marginal dos registros não suspeitos e o método proporcional.

A imputação manual dessas variáveis, para as mulheres com mais de 19 filhos, foi feita, em microcomputador, por dois técnicos, trabalhando as informações reunidas em nível de Brasil. Quando havia informação para o total de filhos e para um dos sexos, o número de filhos para de outro sexo foi imputado pela diferença. Quando havia somente registro para uma das categorias - total de filhos, filhos (as) homens ou mulheres – as demais categorias eram imputadas tomando como base os dados de outra mulher – doador – que apresentasse informações corretas para as três categorias e, também, coincidência com a categoria declarada para o registro a ser imputado.

h) Críticas das Variáveis de Rendimento

Na fase de crítica de consistência dos dados de rendimento das pessoas pesquisadas na amostra do Censo Demográfico 2000, buscou-se verificar para cada tipo de rendimento (trabalho principal, demais trabalhos, de aposentadoria ou pensão, etc.) a ocorrência de registros que, por alguma incorreção, pudessem distorcer os resultados obtidos para cada um deles isoladamente e, consequentemente, para o rendimento total.

Para auxiliar nesta verificação, além da geração de alguns indicadores por município, foram selecionados, por unidade da federação, os valores extremos para cada tipo de rendimento e, para o do trabalho principal, por posição na ocupação e categoria do emprego, associados a outras características das pessoas e dos domicílios de residência, a fim de se ter uma primeira visão de inconsistências entre características. O exame mais aprofundado dos casos de inconsistências entre características por meio de consulta às imagens dos questionários revelou dois tipos de problemas. No primeiro, registros de rendimento de trabalho mostravam-se inconsistentes em decorrência de classificações incorretas das características do trabalho ocorridas na fase de coleta. Foram verificadas classificações de posição na ocupação incompatíveis com as ocupações e atividades econômicas. Para estas situações foram definidos procedimentos de critica automatizada restritos aos registros de trabalho.

Verificaram-se ainda, em muito menor quantidade que as observadas durante os trabalhos de crítica dos resultados do universo, principalmente, situações de reconhecimento como valores, de traços delimitadores das quadrículas desenhadas para registro dos algarismos e, ainda, de sombras decorrentes de registros feitos nos versos de algumas das folhas do questionário. Para estes casos, foram definidos procedimentos automatizados de eliminação dos registros indevidos do rendimento, considerando que, na maior parte dos casos, eram de valores formados por seqüência do dígito 1, combinado ou não com o dígito 7; e, para definir as situações em que deveria haver imputação do rendimento, foram considerados outros registros indicativos de sua existência ou não.

Concluída esta etapa, avaliaram-se, de forma global, os resultados e, de forma pontual, os valores extremos, inclusive com exame das imagens dos questionários. Os valores com fortes inconsistências foram ignorados para serem tratados na etapa de crítica e imputação dos rendimentos.

12.5.2.3 Análise do processo de crítica e imputação

No Censo Demográfico 2000, os procedimentos de crítica e imputação dos dados foram constantemente monitorados a fim de evitar alterações na estrutura da informação. Vários foram os instrumentos utilizados com esse objetivo, como as tabelas (conjunto de indicadores que são calculados na execu-

ção do programa de crítica e imputação), a análise demográfica, os estudos de população e o controle das alterações nas respostas originais constantes do questionário. Este texto trata sobre esse último instrumento.

a) análise dos relatórios do sistema DIA

Os relatórios gerados pelo sistema DIA permitem que sejam realizadas diversas análises, entre elas conhecer o nível de imputação de cada uma das variáveis tratadas no sistema. No caso do questionário da amostra, foram tratadas, através dos diversos aplicativos desenvolvidos no DIA, 22 variáveis do bloco de características do domicílio e 65 do bloco de características dos moradores, considerando-se os diversos temas envolvidos. No total, foram 20 274 412 registros de pessoas e 5 304 711 registros de domicílios.

No bloco de domicílios, a variável com maior índice de imputação (4,15% do total de registros, correspondentes a 220 240 ocorrências) foi a que se refere ao número de cômodos servindo de dormitório. Outras 7 variáveis tiveram imputação entre 1,16% e 1,89% do total de registros e as 14 restantes tiveram imputação em menos de 1,00% dos registros.

No bloco de características de moradores, observou-se a seguinte distribuição nos percentuais de imputação:

Tabela 12.33 - Número de variáveis imputadas, segundo o percentual de imputação

Percentual de imputação	Número de variáveis
Total	**65**
De 0,01 a 1,00	27
De 1,01 a 2,00	27
De 2,01 a 3,50	8
De 3,51 a 5,13	0
De 5,14 a 7,47	3

Fonte: IBGE, Censo Demográfico 2000.

As 3 variáveis com maiores percentuais de imputação (5,14%, 6,09% e 7,47%, correspondentes a 1 042 499, 1 234 566 e 1 515 441 ocorrências, respectivamente) foram as que se referem, pela ordem, à atividade principal em que a pessoa trabalhava na semana de referência, ao município onde trabalha ou estuda – para quem trabalhava ou estudava em município diferente do de residência – e ao curso mais elevado no qual concluiu pelo menos uma série – para quem não freqüentava escola mas já havia freqüentado.

Ao longo deste capítulo, serão detalhados os diversos procedimentos de análise e validação das imputações. Por outro lado, observa-se que 40,0% das variáveis tiveram menos de 1,00% dos registros imputados e outras 40,0% sofreram imputação entre 1,01% e 2,00% dos registros, significando uma boa qualidade geral de preenchimento dos questionários em campo.

Outra análise do processo de crítica e imputação dos Resultados da Amostra decorrentes da aplicação do sistema DIA foi realizada para cada um dos aplicativos da crítica intra-registros, tendo como orientação diretrizes bastante semelhantes às estabelecidas no item 12.3.2.6, que tratou da mesma tarefa para os Resultados do Universo. Os elementos para realização desse trabalho estão

detalhados no Plano de análise da correção automática e elementos de apoio para a análise da formação do Lote siatema DIA, CD 1.02 – questionário da amostra 2002; assim, apontaremos aqui somente os pontos distintos do processo.

No tocante à análise em nível do lote, foram considerados os mesmos indicadores E e F para detecção das situações suspeitas; no entanto, os limites de tolerância estabelecidos foram, respectivamente, 5% e 10%, para os aplicativos de fecundidade – "Fecundidade I, II e III"- e 12% e 50% para os demais aplicativos.

Da mesma forma, a análise em nível de município, considerou também os indicadores E, Ij, Maxi Dj(i), Maxi Aj(i), Tj e Bj , levando em conta, no entanto, os seguintes limites de tolerância, respectivamente:

- indicador E – 12%;

- indicador Ij, – 2% para fecundidade e 3% para os demais aplicativos; e

- indicadores Maxi Dj(i), Maxi Aj(i), Tj e Bj – 2% para os aplicativos de fecundidade e, respectivamente 3%, 3%, 10% e 5% para os demais aplicativos.

b) análise da consistência da imputação e validação dos resultados

É importante fazer uma primeira observação com respeito à Análise do processo de Crítica e Imputação e à dificuldade em separá-la do item validação dos resultados.

Para avaliar os efeitos do processo de imputação na informação obtida, muitas vezes, esta era comparada com a proveniente de outras fontes e com a série histórica com o intuito de descobrir se os indicadores eram compatíveis com o comportamento esperado. Este foi o mesmo recurso utilizado posteriormente na validação dos resultados. Houve grande dificuldade em tratar estes aspectos separadamente já que, muitas vezes, as divergências com a tendência esperada levaram a descobrir problemas no processo analisado.

Portanto, neste item serão abordadas tanto as análises dos processos de crítica e imputação quanto a validação dos resultados obtidos.

Foram realizadas análises para algumas Unidades da Federação e, em alguns casos, para alguns municípios onde era estudado o comportamento de cada quesito do questionário antes e depois da imputação automática. Para isso, foram utilizados a visualização de imagens, a listagem de registros completos das pessoas, contendo as variáveis a serem analisadas, e as matrizes de contingência, conforme modelo abaixo:

Figura 12.5 - Matriz de contingência com valores antes e após o processo de imputação

VPOSTERIOR / VANTERIOR	Total 1	%	1	2	3	...	j	...	n
Total2									
%									
1									
2									
3									
...									
j									
...									
m									

Onde:

VANTERIOR = código do quesito antes da correção automática

VPOSTERIOR = código do quesito após a correção automática

A acumulação de valores na diagonal indica ausência de modificações no processo, portanto, constituíram o alvo do estudo os casos em que houve significativo aumento, ou redução, de frequência de casos observados em alguma categoria fora da diagonal. As distorções, em geral, desapareceram com a revisão e alteração de algumas regras contidas no plano de crítica. Em alguns casos, elas estavam justificadas por se tratar de correção de erros sistemáticos.

As listagens dos registros e as matrizes de contingência foram obtidas através da utilização do REDATAM+G4 (REcuperação de DAdos para Áreas pequenas por Microcomputador, 4ª Geração), um programa computacional desenvolvido pelo Centro Latino-americano e Caribenho de Demografia – CELADE. Com este objetivo, foi feita a junção dos arquivos antes e depois da imputação.

Faixa de idade

A faixa de idade é uma variável auxiliar, criada para ajudar na imputação de dados e no controle da consistência entre as variáveis imputadas e as demais variáveis relacionadas.

Inicialmente, o cálculo da faixa de idade levou em consideração somente as respostas obtidas em alguns quesitos do bloco de nupcialidade ("Vive em companhia de cônjuge ou companheiro?", "Qual é (era) a natureza da última união?" e "Qual o seu estado civil?"); de mão-de-obra ("Na semana de 23 a 29 de julho trabalhou em atividade de cultivo, extração vegetal, criação de animais ou pesca, destinados à alimentação de pessoas moradoras no domicílio?", "Quantos trabalhos tinha na semana de 23 a 29 de julho?", "Nesse trabalho era:", "No período de 30 de junho a 29 de julho de 2000, tomou alguma providência para conseguir algum trabalho?" e "Em julho de 2000, era aposentado de instituto de previdência oficial?"); de fecundidade ("Quantos(as) filhos(as) nascidos(as) vivos(as) teve até 31 de julho de 2000?" e "Quantos(as) filhos(as) nascidos(as) mortos(as) teve até 31 de julho de 2000?"); de relação com a pessoa responsável pelo domicílio ou de relação com a pessoa responsável pela família.

Foi estabelecida faixa 1 (0 a 9 anos de idade), quando não havia qualquer resposta aos quesitos selecionados de nupcialidade, mão-de-obra e fecundidade, faixa 2 (10 anos ou mais de idade) para os casos em que havia resposta válida em todos os quesitos selecionados de nupcialidade, mão-de-obra e fecundidade, mas não se tratava de pai/mãe ou sogro/sogra do responsável pelo domicílio ou pela família, e faixa 3 (mais de 20 anos) quando, além de haver resposta válida em todos os quesitos selecionados de nupcialidade, mão-de-obra e fecundidade, tratava-se de pai/mãe ou sogro/sogra do responsável pela família ou pelo domicílio.

Embora as perguntas referentes aos quesitos de nupcialidade, fecundidade e mão-de-obra só se apliquem a pessoas com 10 anos ou mais de idade, em alguns casos havia respostas válidas para pessoas com idade entre 0 e 9 anos. Tais respostas ocorreram, principalmente, para o bloco de nupcialidade, sendo que estas eram coerentes com a idade das pessoas para as quais foram obser-

vadas. Em outras palavras, as crianças declararam nunca terem vivido em companhia de cônjuge ou companheiro, assim como declararam serem solteiras. Nestes casos, a pessoa foi alocada nas faixas 2 ou 3 e sua idade trocada posteriormente, quando compatibilizada com a "faixa de idade".

A troca de idade acarretou uma mudança na distribuição de idade, com diminuição do número de crianças de 0 a 9 anos e conseqüente aumento das pessoas com outras idades.

Para a correção do problema, a idade da pessoa foi introduzida no processo de construção da "faixa de idade". Assim, a faixa passou a ser calculada, levando em consideração a idade, sempre que a idade calculada a partir do mês e do ano de nascimento (V4075) fosse igual[5] à idade declarada (V4062). No caso de não haver informação de mês e ano de nascimento ou mesmo de haver diferença entre as idades declarada e calculada, a faixa foi determinada pelos quesitos de nupcialidade, mão-de-obra, fecundidade e relação com a pessoa responsável pelo domicílio ou pela família, conforme procedimento anteriormente adotado.

Declaração de Idade

Apesar da idade calculada através de mês e ano de nascimento ser a informação mais robusta entre as investigadas, houve preocupação em melhorar a sua qualidade. Para tanto, além das perguntas tradicionais do mês e ano de nascimento – "Qual é o mês e o ano do seu nascimento?", assim como da idade presumida – "Qual é a sua idade presumida?", no Censo Demográfico 2000 foi introduzida a idade declarada das pessoas, através da pergunta "Qual era a sua idade em 31 de julho de 2000?".

Inicialmente, o cálculo da idade foi feito a partir das informações já tradicionais, mês e ano de nascimento ou idade presumida. Contudo, a nova informação foi fundamental, pois possibilitou uma análise de compatibilidade entre a idade declarada pela pessoa e a idade calculada, utilizando-se mês e ano de nascimento ou idade presumida.

Algumas divergências foram encontradas, e uma análise mais detalhada foi feita para as pessoas com cem anos ou mais de idade, grupo em que o número de casos com diferença era mais significativo. Em todo o Brasil, foram encontrados na amostra, sem expansão, 1.416 casos com divergência entre os dois tipos de idade, no entanto, havia ao todo 2.544 pessoas com 100 anos ou mais de idade.

Através do sistema de consulta de imagens dos questionários, verificou-se ter havido, na maioria dos casos, erro sistemático na grafia da centena do ano de nascimento, no momento da coleta. Nestes casos, as diferenças entre a idade declarada e a idade calculada pelo mês e ano de nascimento eram, predominantemente, de 100 anos[6]. Verificou-se, ainda, que a década de nascimento das pessoas cujas idades apresentaram esta diferença era predominantemente a de 1980, assim ao invés de grafar 1989 o recenseador teria grafado 1889, por exemplo.

Para ter certeza de que a idade correta era a declarada e não a obtida pelo mês e o ano de nascimento, foram analisados o sexo, a relação com o responsável pelo domicílio e alfabetização. A razão de sexos encontrada(1,13) não é típica entre os mais idosos. Além disso, em 82% dos casos as pessoas eram filhas do responsável pelo domicílio e em 92% sabiam ler ou escrever. Portanto, indicadores mais compatíveis com jovens, conforme indicava a idade declarada.

---

[5] Considerou-se igual sempre que o módulo da diferença era 0 ou 1.

[6] Considerou-se diferença de 100 anos sempre que esta estava entre 99 e 101.

Para estes casos, prevaleceu a idade declarada, ou seja, a idade considerada passou a ser a idade declarada sempre que a idade calculada, a partir do mês e ano de nascimento, era maior que 100 anos e a diferença entre ambas era de 100 anos. Adicionalmente, a idade declarada passou a ser incorporada também para o cálculo da variável auxiliar faixa de idade.

Domicílio

Em relação aos dados preliminares do censo demográfico, foram realizadas análises dos resultados nas etapas de crítica referentes ao NIM, Pré-Dia e Pós-Dia. Inicialmente, as análises contemplaram unidades da federação selecionadas, como o Rio de Janeiro, Bahia e Espírito Santo. Nestes casos, os dados se mostraram coerentes na sua estrutura interna e em comparação aos resultados obtidos no Censo Demográfico 1991. Em um segundo momento, a análise foi ampliada para os demais estados do País com resultados semelhantes aos encontrados na etapa anterior.

Família

Foram analisadas as imputações das relações de parentesco realizadas através de procedimentos automáticos de imputação, utilizando o sistema New Imputation Methodology – NIM.

Os procedimentos utilizados para validação dos resultados foram obtidos a partir da comparação com os resultados de outras pesquisas domiciliares anteriores que continham as mesmas questões, como por exemplo PNADs da década de 1990 e o Censo Demográfico 1991, através de tabulações que apresentavam a distribuição dos tipos de família. Os resultados mostraram que as imputações não alteraram as distribuições das variáveis envolvidas.

Filtros

Tendo em vista que a captura dos dados foi feita através de processo de leitura e reconhecimento ótico de caracteres (ICR) e marcas, foi necessário criar procedimentos para minimizar os erros provenientes de sombras, sujeiras e outras alterações que pudessem modificar as informações advindas dos questionários. Com essa finalidade, foram estabelecidos limites de valores para as respostas obtidas em alguns quesitos, os quais foram considerados razoáveis, tomando-se em conta os valores esperados em cada uma das variáveis envolvidas. Para o controle dos valores que ultrapassassem os limites estabelecidos, foi introduzida uma rotina no programa de reconhecimento de caracteres e marcas cuja função foi recomendar verificação visual dos campos onde foram encontrados tais valores. O processo de verificação teve a finalidade de confirmar os valores lidos pelo *scanner* e não de criticá-los, ou seja, uma vez confirmado o valor através da imagem do questionário, mesmo sendo improvável, o verificador deveria confirmá-lo, pois a crítica e análise de consistência seriam feitas em etapa posterior.

Em razão de terem sido encontrados, já na fase de análise, valores improváveis em algumas variáveis, suspeitou-se que a rotina introduzida no programa de reconhecimento de caracteres e marcas havia falhado, não obstante tivesse funcionado em uma massa de teste feita na ocasião de sua introdução. Para testar o funcionamento da rotina, foi confeccionado um questionário contendo valores fora dos respectivos limites para os campos "Quantos banheiros

existem neste domicílio?", "Qual o mês e ano do seu nascimento?", "Qual era a sua idade em 31 de julho de 2000?", "Qual é a sua idade presumida?" e "Qual foi o seu rendimento bruto do mês de julho de 2000?". No teste, os campos com problemas não foram selecionados para a verificação visual, portanto, foi confirmado que a rotina não estava funcionando.

Considerando-se que este problema afetou apenas a etapa de reconhecimento dos caracteres e marcas e que as ferramentas utilizadas permitiam o reprocessamento com alguma agilidade, foi possível executar novamente a etapa reconhecimento de caracteres e fazer a verificação visual nos casos previstos na rotina.

Como já havia sido feita uma análise exploratória dos dados da amostra, aproveitou-se este momento para solicitar a inclusão de novos valores limites assim como novas variáveis com respectivos limites de valores a serem aceitos sem verificação visual.

Migração

A maioria dos problemas detectados na informação sobre movimentos migratórios, proveniente do campo, foi solucionada a partir de mudanças realizadas no processo de imputação pelo sistema DIA, com exceção do problema referente ao entendimento do quesito 4.23, no qual se investigava "a Unidade da Federação ou país estrangeiro de residência anterior". No processo de análise da informação proveniente do campo, observou-se, em alguns casos, um não entendimento deste quesito. Na realidade, o objeto investigado era a Unidade da Federação na qual o indivíduo residia antes de mudar-se para a Unidade da Federação em que foi recenseado. Esta formulação propiciou um entendimento equivocado do objeto em estudo, na medida em que muitos indivíduos declararam morar há menos de dez anos na Unidade da Federação em que foram recenseados e responderam a própria unidade como residência anterior. Uma das possíveis explicações para esta observação, tem como origem uma interpretação onde confundiu-se "Unidade da Federação de residência anterior" - conforme o texto da pergunta no questionário – com "Unidade da Federação da residência anterior". Assim, muitos dos indivíduos que efetuaram movimentos entre municípios do estado em que foram recenseados, depois de terem ali chegado, provenientes de outra Unidade da Federação, consideraram este último movimento e declararam como Unidade da Federação de residência anterior a própria unidade. Estes casos foram incluídos na categoria "ignorado".

Em virtude da expressiva quantidade de pessoas, cujo tempo ininterrupto de residência na Unidade da Federação era menor que a idade declarada e, consequentemente, teriam que ter declarado outra unidade diferente daquela em que foram investigados, mas não o fizeram, ficou decidido que a informação desses indivíduos seria considerada "ignorada". A justificativa para este fato é proveniente da forma do processo de imputação deste quesito, onde todas as Unidades da Federação e países estrangeiros tinham possibilidade de serem selecionados, segundo a distribuição dentro do lote, o que podia distorcer os fluxos migratórios.

Este fato foi observado em todas as Unidades da Federação, como mostra a Tabela 12.34. Com relação à população não-natural do estado, dos 8.691.756 que responderam ao quesito, 313.590 foram considerados ignorados, por terem respondido o mesmo estado em que foram recenseados, representando 3,6% do total. Os maiores percentuais foram encontrados nos estados do Mato

Grosso (7,6%), Pará (5,2%), São Paulo (4,6%) e Rondônia (4,6%). O menor percentual foi encontrado no Distrito Federal (0,1%), fato esperado, pois só existe um município, o próprio Distrito Federal.

Tabela 12.34 - Pessoas com menos de 10 anos ininterruptos de residência que responderam ao quesito de residência anterior cujas declarações foram consideradas ignoradas - 2000

Unidades da Federação atual	Não naturais com menos de 10 anos ininterruptos de residência		
	Total	Ignorado	% de ignorado
Brasil	**8 691 756**	**313 590**	**3,6**
Rondônia	173 263	7 981	4,6
Acre	23 967	332	1,4
Amazonas	144 991	2 696	1,9
Roraima	83 765	624	0,7
Pará	355 198	18 557	5,2
Amapá	89 055	1 487	1,7
Tocantins	170 058	6 114	3,6
Maranhão	128 687	3 835	3,0
Piauí	95 809	1 258	1,3
Ceará	165 289	2 907	1,8
Rio Grande do Norte	94 392	1 494	1,6
Paraíba	108 909	2 196	2,0
Pernambuco	194 921	5 183	2,7
Alagoas	74 896	2 287	3,1
Sergipe	79 868	1 649	2,1
Bahia	296 706	8 657	2,9
Minas Gerais	548 244	13 649	2,5
Espírito Santo	221 429	4 808	2,2
Rio de Janeiro	561 315	22 891	4,1
São Paulo	2 638 297	122 676	4,6
Paraná	426 257	16 185	3,8
Santa Catarina	327 143	7 499	2,3
Rio Grande do Sul	149 593	2 988	2,0
Mato Grosso do Sul	176 171	5 824	3,3
Mato Grosso	362 108	27 385	7,6
Goiás	598 356	22 149	3,7
Distrito Federal	403 070	282	0,1

Fonte: IBGE, Censo Demográfico 2000.

Educação

Em primeiro lugar, serão apresentados os limites etários estabelecidos para a imputação de dados relativos à freqüência escolar para, em seguida, serem relatadas as inconsistências verificadas após a aplicação do sistema DIA, assim como os procedimentos adotados para correção.

Limites etários estabelecidos para a imputação de dados nos quesitos "Qual é o curso que frequenta?" e "Qual é a série que frequenta?".

Os limites inferiores e superiores de idade para a freqüência escolar em cada nível ou modalidade de ensino foram definidos, levando-se em conta o disposto na Lei de Diretrizes e Bases da Educação – LDB – além de discutidos com técnicos do Ministério da Educação e Cultura – MEC, responsáveis pela elaboração do Censo Escolar. Sendo assim, as idades consideradas adequadas para cursar os diversos níveis e modalidades de ensino serviram de referência para a construção de tais limites, que foram ampliados, levando-se em consideração a realidade observada no campo.

Educação Infantil (Creche e Pré-escola)

Foram consideradas as idades pontuais entre 0 e 11 anos, sendo de 0 a 4 anos para freqüência à creche e 3 a 11 anos para pré-escola.

Classe de alfabetização

Para o curso regular, foram consideradas as idades de 4 a 11 anos. Para a classe de Alfabetização de Adultos, foram consideradas as idades a partir de 13 anos.

Ensino Fundamental

Para a freqüência neste nível de ensino, foram considerados os seguintes limites etários: regular seriado ou não seriado, mínimo 5 anos de idade; e supletivo de 1º grau, mínimo 13 anos de idade.

Ensino Médio

Para a freqüência neste nível de ensino, foram considerados os seguintes limites etários: regular seriado ou não seriado, mínimo 13 anos de idade; e supletivo de 2º grau, mínimo 17 anos de idade.

Educação Superior

Foram considerados os seguintes limites etários: graduação, mínimo 16 anos de idade e mestrado ou doutorado, mínimo 20 anos de idade.

Análise dos resultados

Inicialmente, foi feita uma análise de caráter exploratório a partir de tabelas de contingência e da construção de indicadores. Nesta etapa, observou-se que as informações sobre o analfabetismo estavam compatíveis com o esperado. Paralelamente, foi realizado um estudo da estrutura etária, através do qual detectou-se que a aplicação do sistema DIA acarretou alterações nas freqüências das idades de 0 a 2, 5 e 6 anos, pois promoveu alteração nas idades para que estas ficassem compatíveis com os níveis de ensino freqüentados.

Para corrigir as mudanças na estrutura etária, foram elaboradas algumas regras de imputação determinística. Na elaboração, foram consideradas as informações adicionais obtidas com a análise detalhada dos dados e também com a troca de experiência com os técnicos do MEC. As regras aplicadas foram as seguintes:

primeira mudança no nível freqüentado – pessoas com 5 ou 6 anos de idade, com declaração de freqüência à creche, passaram a freqüência da pré-escola; e

segunda mudança no nível freqüentado – crianças com 0, 1 e 2 anos de idade, com declaração de freqüência à pré-escola, passaram a freqüência da creche.

A análise das variáveis "Qual é o curso mais elevado que freqüentou, no qual concluiu pelo menos uma série?", "Qual é a última série concluída com aprovação?" e "Concluiu o curso no qual estudou?" foi feita, inicialmente, com os dados do Espírito Santo, Rio de Janeiro e Bahia.

Detectou-se uma alteração sistemática na distribuição de freqüência da variável "Qual é o curso mais elevado que freqüentou, no qual concluiu pelo menos uma série?", após a aplicação do sistema DIA. Esta alteração foi verificada sempre que a pessoa, com idade entre 17 e 37 anos e que concluiu o curso no qual estudou, possuía declaração de conclusão do antigo primário, ensino fundamental ou 1ª grau e a última série concluída com aprovação era 4ª e, em alguns casos, 3ª. Nesses casos, a informação do curso mais elevado que freqüentou, no qual concluiu pelo menos uma série, foi alterada para ensino médio ou 2º grau, mais compatível com as séries informadas e o curso concluído.

Em razão desta alteração, aumentou o número de pessoas com 11 anos de estudo e, conseqüentemente, diminuiu o número das que tinham 4 anos. No entanto, a média de anos de estudo para a população de 10 anos ou mais de idade apresentou uma alteração pequena, em torno de 0,2 pontos percentuais.

Em consulta feita ao MEC, foi obtida a informação de que, até recentemente, as escolas forneciam certificado de conclusão do primeiro segmento do ensino fundamental às pessoas que concluíam este nível através de curso supletivo. Essa informação nos levou a considerar que os recenseadores podiam ter entendido que as pessoas concluíram o curso fundamental e não apenas o primeiro segmento do fundamental, como seria o correto.

Para corrigir as alterações detectadas, criaram-se as seguintes críticas:

1. se a pessoa tinha mais de 16 e menos de 37 anos e concluiu o antigo primário em 3 ou 4 séries, então o curso mais elevado que freqüentou, no qual concluiu pelo menos uma série, passava para ensino fundamental ou 1º grau incompleto. Este procedimento também foi aplicado aos casos em que o quesito "Concluiu o curso no qual estudou?" estava em branco; e.

2. se a pessoa tinha mais de 16 anos e concluiu o ensino fundamental ou 1º grau em 3 ou 4 séries, então se declarou sim no quesito "Concluiu o curso no qual estudou?" passava para não concluiu. Este procedimento também foi aplicado ao caso em que o quesito "Concluiu o curso no qual estudou?" estava em branco.

Trabalho e Rendimento

Durante a fase de crítica e imputação dos dados da parte de trabalho do Questionário da Amostra do Censo Demográfico 2000, foram realizadas análises de consistência para cada característica investigada, considerando os resultados de cada quesito e sua associação com outras informações do próprio questionário. Estas análises revelaram a existência de algumas falhas sistemáticas, oriundas da fase de coleta, que o procedimento geral de crítica e imputação não poderia solucionar. Assim sendo, foram adotados procedimentos específicos e de imputação determinística para resolver estas situações. Constatou-se, ainda, que para determinados segmentos da população, a imputação de quesitos omitidos causava distorção nos resultados de algumas características. Também para solucionar estas situações, foram adotados procedimentos de imputação determinística.

Ao final de todo o processo de apuração, foi desenvolvido o estudo, apresentado a seguir, sobre a imputação das informações do tema trabalho e rendimento.

A análise da imputação das variáveis foi iniciada com a identificação das informações dos indivíduos que, originalmente, não tinham qualquer resposta aos quesitos do tema trabalho e rendimento (todos eles em branco). Para todo o Brasil, ocorreram 24 268 desses casos, que representam 0,1% do total de pessoas com 10 anos ou mais de idade da amostra (16 194 309).

Nesse grupo de casos que estavam originalmente em branco, 50 foram imputados como ocupados, dos quais 46 foram classificados como trabalhadores domésticos (44 deles com carteira de trabalho assinada e 2 sem carteira assinada) e 4 como empregados com carteira de trabalho assinada. Outros 2 117 desses casos foram imputados como pessoas desocupadas. Os demais (22 101 casos) foram imputados, portanto, como pessoas não-economicamente ativas.

No grupo de pessoas imputadas como "não-economicamente ativas" (22 101 casos), a maior parte estava na faixa de 10 a 14 anos de idade (71,2%), sendo que 15 402 pessoas não sofreram imputação na variável V4752 "idade calculada em anos" (69,7%), conforme tabela a seguir.

Tabela 12.35 - Pessoas que tinham as variáveis de trabalho e rendimento originalmente em branco e foram imputadas como não economicamente ativas, por condição de imputação na variável V4752, segundo os grupos de idade

Grupos de idade	Pessoas que tinham as variáveis de trabalho e rendimento originalmente em branco e foram imputadas como não economicamente ativas		
	Total	Condição de imputação na variável V4752	
		Sem imputação	Com imputação
	Valores absolutos		
Total	**22 101**	**21 306**	**795**
10 a 14 anos	15 732	15 402	330
15 a 24 anos	3 137	2 956	181
25 anos ou mais	3 232	2 948	284
	Valores relativos		
Total	**100,0**	**96,4**	**3,6**
10 a 14 anos	71,2	69,7	1,5
15 a 24 anos	14,2	13,4	0,8
25 anos ou mais	14,6	13,3	1,3

Fonte: IBGE, Censo Demográfico 2000.

Posição na ocupação

O estudo dos temas trabalho e rendimento foi iniciado com a variável V0447, que define a posição na ocupação. A partir do levantamento das proporções de informações imputadas no processo de apuração, identificou-se a necessidade de aprofundar a caracterização de algumas posições tais como trabalhadores domésticos e aprendizes ou estagiários.

Trabalhadores Domésticos

No estudo da variável "posição na ocupação", verificou-se que as proporções de imputações de informações nos códigos 1 (trabalhador doméstico com carteira de trabalho assinada) e 2 (trabalhador doméstico sem carteira de trabalho assinada) da variável V0447, ao final do processo de apuração, foram extremamente elevadas, como pode ser constatado na Tabela 12.36 a seguir:

Tabela 12.36 - Pessoas registradas na variável V0447, com e sem imputação, segundo as categorias da variável

Categorias da variável	Pessoas registradas na variável V0447					
	Valores absolutos			Valores relativos (%)		
	Total	Sem imputação	Com imputação	Total	Sem imputação	Com imputação
Total	**20 274 412**	**19 910 122**	**364 290**	**100,0**	**98,2**	**1,8**
Trabalhador doméstico com carteira de trabalho assinada	164 674	100 895	63 779	100,0	61,3	38,7
Trabalhador doméstico sem carteira de trabalho assinada	414 090	286 083	128 007	100,0	69,1	30,9
Empregado com carteira de trabalho assinada	2 470 074	2 411850	58 224	100,0	97,6	2,4
Empregado sem carteira de trabalho assinada	1 930 353	1 901 589	28 764	100,0	98,5	1,5
Empregador	209 478	207 903	1 575	100,0	99,2	0,8
Conta própria	1 858 789	1 806 211	52 578	100,0	97,2	2,8
Aprendiz ou estagiário sem remuneração	32 256	30 446	1 810	100,0	94,4	5,6
Não remunerado em ajuda a membro do domicílio	393 503	391 112	2 391	100,0	99,4	0,6
Trabalhador na produção para o próprio consumo	309 682	308 196	1 486	100,0	99,5	0,5
Branco	12 491 513	12 465 837	25 676	100,0	99,8	0,2

Fonte: IBGE, Censo Demográfico 2000.

A fim de permitir um melhor entendimento e uma avaliação mais adequada dos fatores que levaram ao elevado número de imputações nos itens 1 e 2 dessa variável, foram feitos, inicialmente, outros cruzamentos de variáveis. Para orientar esses novos cruzamentos, levou-se em consideração a informação sobre ocupação das pessoas selecionadas e a atividade econômica do empreendimento em que trabalhavam.

Para as pessoas classificadas como trabalhadoras domésticas com imputação na variável V0447, foi verificada a ocorrência de imputações nas variáveis que identificavam a atividade do empreendimento em que trabalhavam (V4462) e a ocupação da pessoa (V4452). O resultado desse cruzamento revelou que:

1 - Para o conjunto de 63 779 pessoas que, ao final do processo de apuração, foram imputadas como trabalhadoras domésticas com carteira de trabalho assinada, 52 555 não apresentavam marca de imputação na variável identificadora da ocupação (V4452) e nem na variável identificadora de atividade econômica do empreendimento (V4462). Assim sendo, 82,4% das pessoas com imputação no item 1 da variável V0447 foram assim classificadas em função da sua ocupação e da atividade econômica; portanto, estavam claramente identificadas como trabalhadoras domésticas. Constataram-se 6 casos de pessoas sem imputação na variável de atividade (que identifica a atividade econômica do serviço doméstico em um código especifico) e com imputação na variável de ocupação e, ainda, 9 242 casos de pessoas sem imputação na variável de ocupação (em que, do conjunto de ocupações admissíveis para a atividade econômica do serviço doméstico, uma parcela é exclusiva dessa atividade) e com imputação na variável de atividade econômica. Em 1 976 casos (3,1% dos casos em que houve imputação no item 1 da variável V0447), as variáveis de ocupação e atividade também foram imputadas.

Tabela 12.37 - Pessoas registradas com imputação no item 1 - 'trabalhador doméstico com carteira de trabalho assinada' da variável V0447, por condição de imputação na variável V4462 - atividade, segundo a condição de imputação na variável V4452 - ocupação

Condição de imputação na variável V4452 - ocupação	Pessoas registradas com imputação no item 1 - trabalhor doméstico com carteira de trabalho assinada da variável V0447		
	Total	Condição de imputação na variável V4462 - atividade	
		Sem imputação	Com imputação
	Valores absolutos		
Total	**63 779**	**52 561**	**11 218**
Sem imputação	61 797	52 555	9 242
Com imputação	1 982	6	1 976
	Valores relativos		
Total	**100,0**	**82,4**	**17,6**
Sem imputação	96,9	82,4	14,5
Com imputação	3,1	0,0	3,1

Fonte: IBGE, Censo Demográfico 2000.

2 - Para o conjunto de 128 007 pessoas que, ao final do processo de apuração, foram imputadas como trabalhadoras domésticas sem carteira de trabalho assinada, 112 675 não apresentavam marca de imputação na variável identificadora da ocupação (V4452) e nem na variável identificadora da atividade econômica (V4462). Assim sendo, 88,0% das pessoas com imputação no item 2 da variável V0447 em função da sua ocupação e atividade estavam claramente identificadas como trabalhadoras domésticas. Constataram-se, também, 20 casos de pessoas sem imputação na variável de atividade (que identifica a atividade econômica do serviço doméstico em um código especifico) e com imputação na variável de ocupação e, ainda, 14 324 casos de pessoas sem imputação na variável de ocupação (em que o conjunto de ocupações admissíveis para a atividade econômica do serviço doméstico, uma parcela é específica unicamente dessa atividade). Em 988 casos (0,8% dos casos em que houve imputação no item 2 da variável V0447), as variáveis de ocupação e atividade também foram imputadas.

Tabela 12.38 - Pessoas registradas com imputação no item 2 - 'trabalhador doméstico sem carteira de trabalho assinada' da variável V0447, por condição de imputação na variável V4462 - atividade, segundo a condição de imputação na variável V4452 - ocupação

Condição de imputação na variável V4452 - ocupação	Pessoas registradas com imputação no item 2 - trabalhor doméstico sem carteira de trabalho assinada da variável V0447		
	Total	Condição de imputação na variável V4462 - atividade	
		Sem imputação	Com imputação
	Valores absolutos		
Total	**128 007**	**112 695**	**15 312**
Sem imputação	126 999	112 675	14 324
Com imputação	1 008	20	988
	Valores relativos		
Total	**100,0**	**88,0**	**12,0**
Sem imputação	99,2	88,0	11,2
Com imputação	0,8	0,0	0,8

Fonte: IBGE, Censo Demográfico 2000.

Com esses primeiros cruzamentos, justifica-se a maior parte das informações imputadas para pessoas classificadas nas posições na ocupação como trabalhadoras domésticas com e sem carteira de trabalho assinada (respectivamente, itens 1 e 2 da variável V0447) a partir das informações sobre ocupação e atividade econômica correspondentes.

Entre outras investigações, ainda foi verificada a parcela em que se têm indicativos sobre carteira de trabalho assinada. Do conjunto de pessoas classificadas como trabalhadoras domésticas com carteira de trabalho assinada, com imputação dessa característica, 78,1% (49 812) tinham sido originalmente classificadas como empregadas com carteira de trabalho assinada, sendo que em 66,5% dos casos (42 405) não houve imputação de ocupação nem de atividade. Assim, a identificação original de pessoa com registro em carteira de trabalho foi mantida após a imputação de posição na ocupação na categoria de trabalhadores domésticos.

Tabela 12.39 - Pessoas registradas com imputação no item 1 - 'trabalhador doméstico com carteira de trabalho assinada' da variável V0447, por condição de imputação na variável V4462 - atividade, segundo a condição de imputação na variável V4452 - ocupação

Condição de imputação na variável V4452 - ocupação	Pessoas registradas com imputação no item 1 - trabalhor domético com carteira de trabalho assinada da variável V0447, originalmente classificadas como empregador com carteira trabalho assinada		
	Total	Condição de imputação na variável V4462 - atividade	
		Sem imputação	Com imputação
	Valores absolutos		
Total	**49 812**	**42 406**	**7 406**
Sem imputação	49 792	42 405	7 387
Com imputação	20	1	19
	Valores relativos		
Total	**100,0**	**85,1**	**14,9**
Sem imputação	100,0	85,1	14,8
Com imputação	0	0,0	0,0

Fonte: IBGE, Censo Demográfico 2000.

Essa informação foi também importante na imputação das pessoas classificadas como trabalhadoras domésticas sem carteira de trabalho assinada. Dessas pessoas, 81,8% (104 726) tinham sido originalmente classificadas como empregadas sem carteira de trabalho assinada, sendo que para 73,0% delas (93 397) não houve imputação de ocupação nem de atividade.

Tabela 12.40 - Pessoas registradas com imputação no item 2 - 'trabalhador doméstico sem carteira de trabalho assinada' da variável V0447, por condição de imputação na variável V4462 - atividade, segundo a condição de imputação na variável V4452 - ocupação

Condição de imputação na variável V4452 - ocupação	Pessoas registradas com imputação no item 2 - trabalhor domético sem carteira de trabalho assinada da variável V0447 originalmente classificadas como empregador sem carteira trabalho assinada		
	Total	Condição de imputação na variável V4462 - atividade	
		Sem imputação	Com imputação
	Valores absolutos		
Total	**104 726**	**93 398**	**11 328**
Sem imputação	104 704	93 397	11 307
Com imputação	22	1	21
	Valores relativos		
Total	**100,0**	**89,2**	**10,8**
Sem imputação	100,0	89,2	10,8
Com imputação	0,0	0,0	0,0

Fonte: IBGE, Censo Demográfico 2000.

Portanto, as informações sobre a ocupação da pessoa e a atividade econômica do empreendimento em que trabalhava, associadas a informações sobre registro em carteira de trabalho, explicam a maior parte das imputações da posição na ocupação de trabalhadores domésticos com e sem carteira de trabalho assinada.

Por outro lado, no grupo de pessoas com imputação da posição na ocupação como trabalhadoras domésticas com carteira de trabalho assinada, 1 886 não registraram originalmente informações nas variáveis de posição na ocupação, ocupação nem atividade do empreendimento em que trabalhavam, o que corresponde a 3,0% desse grupo ao final do processo de apuração e crítica. Ou seja, as variáveis V0447, V4452 e V4462 estavam originalmente em branco. Para as pessoas com imputação da posição na ocupação como trabalhadoras domésticas sem carteira de trabalho assinada, 199 delas (ou 0,2%) não possuíam informações originais nessas variáveis. Nesses casos, outras informações foram consideradas no processo de crítica para a definição da posição na ocupação.

Aprendizes ou estagiários sem remuneração

Para as pessoas classificadas como aprendizes ou estagiárias sem remuneração, a proporção de imputação do código de posição na ocupação foi de 5,6%, que corresponde a 1 810 indivíduos.

Considerando esse grupo de pessoas com informação imputada de posição na ocupação "aprendiz ou estagiário sem remuneração", 1 696 pessoas declararam que ajudaram, sem remuneração, no trabalho exercido por pessoa conta própria ou empregadora, moradora do domicílio, ou como aprendiz ou estagiário (V0441), conforme descrito na tabela a seguir.

Tabela 12.41 - Pessoas registradas com imputação no item 7 - Aprendizes ou estagiários sem remuneração da variável V0447, por condição de imputação na variável V0441,segundo a variável V0441

Variável V0441	Pessoas registradas com imputação no item 7 - Aprendizes ou estagiários sem remuneração da variável V0447		
	Total	Condição de imputação na variável V0441	
		Sem imputação	Com imputação
	Valores absolutos		
Total	**1 810**	**1 729**	**81**
Sem imputação	1 776	1 696	80
Com imputação	34	33	1
	Valores relativos		
Total	**100,0**	**95,5**	**4,5**
Sem imputação	98,1	93,7	4,4
Com imputação	1,9	1,8	0,1

Fonte: IBGE, Censo Demográfico 2000.

O registro do código 1 da variável V0441 poderia levar à imputação como "aprendiz ou estagiário sem remuneração" ou como "trabalhador não remunerado em ajuda a membro do domicílio"; entretanto, proporcionalmente, houve mais imputação no código 7 (aprendiz ou estagiário sem remuneração) do que no código 8 (trabalhador não remunerado em ajuda a membro do domicílio).

Dois ou mais trabalhos

No que diz respeito aos indivíduos com mais de um trabalho, ocorreram 15 180 casos de imputação de informações, que representaram 5,7% do total de indivíduos nesta situação. Os indivíduos que possuíam informações originais, sem imputação, em pelo menos uma das variáveis horas trabalhadas nos demais trabalhos (V0454), recebiam apenas em benefícios nos demais trabalhos (V4521 igual à opção 1) ou receberam algum valor de rendimento dos demais trabalhos (V4522) eram 11 477, correspondendo a 75,6% dos casos de imputação de dois ou mais trabalhos (V0444).

Observando cada um desses grupos separadamente, 10 425 pessoas informaram horas trabalhadas nos demais trabalhos, 7 325 informaram valor de rendimento nos demais trabalhos e 841 informaram receber somente em benefícios nos demais trabalhos.

Por outro lado, no caso da variável V0454 "horas trabalhadas nos demais trabalhos na semana", 11,6% dos casos (correspondentes a 31 068 pessoas) com horas diferentes de "zero" foram imputados. Desse grupo, 26 313 pessoas (84,7%) declararam originalmente possuir dois ou mais trabalhos na semana de referência (variável V0444 sem imputação).

Rendimentos

A variável indicadora de imputação de que a pessoa não possuía rendimento do trabalho principal (V4511) teve proporção significativa de valores imputados (23,8%, que representam 101 236 pessoas). Nesse grupo, 89,2% das pessoas informaram originalmente que eram trabalhadores não remunerados em ajuda a membro da unidade familiar e outros 10,1% que eram aprendizes ou estagiários sem remuneração. Portanto, 99,2% (ou 100 465 pessoas) tinham informações originais que justificavam a imputação de que não tinham rendimento do trabalho principal.

Tabela 12.42 - Pessoas registradas com imputação no item 0 da variável V4511, por condição de imputação na variável V0447, segundo a variável V0447

Variável V0447	Pessoas registradas com imputação no item 0 (1)		
	Total	Condição de imputação na variável V0447	
		Sem imputação	Com imputação
	Valores absolutos		
Total	**101 236**	**100 465**	**771**
Aprendiz ou estagiário sem remuneração	10 618	10 207	411
Não remunerado em ajuda a membro do domicílio	90 618	90 258	360
	Valores relativos		
Total	**100,0**	**99,2**	**0,8**
Aprendiz ou estagiário sem remuneração	10,5	10,1	0,4
Não remunerado em ajuda a membro do domicílio	89,5	89,2	0,4

Fonte: IBGE, Censo Demográfico 2000.
(1) Não tem rendimento no trabalho principal.

Já a imputação de informações na variável V4521 "não tem rendimento nos demais trabalhos", apesar de pequena em relação ao total de informantes (3,9%), representa 280 193 pessoas. Desse grupo, 89,7% afirmaram originalmente que tinham apenas um trabalho na semana de referência (V0444), o que justifica, portanto, a maior parte da imputação na V4521 (251 313 pessoas).

Tabela 12.43 - Pessoas registradas com imputação no item 0 da variável V4521, por condição de imputação na variável V0444, segundo a variável V0444

Variável V0444	Pessoas registradas com imputação no item 0 (1)		
	Total	Condição de imputação na variável V0444	
		Sem imputação	Com imputação
	Valores absolutos		
Total	**280 193**	**255 747**	**24 446**
Aprendiz ou estagiário sem remuneração	274 895	251 313	23 582
Não remunerado em ajuda a membro do domicílio	5 298	4 434	864
	Valores relativos		
Total	**100,0**	**91,3**	**8,7**
Aprendiz ou estagiário sem remuneração	98,1	89,7	8,4
Não remunerado em ajuda a membro do domicílio	1,9	1,6	0,3

Fonte: IBGE, Censo Demográfico 2000.
(1) Não tem rendimento nos demais trabalhos.

As proporções de pessoas com informações imputadas de rendimentos não oriundos de trabalho foram baixas, sempre inferiores a 0,8% do total, conforme tabela a seguir:

Tabela 12.44 - Pessoas registradas nas variáveis de rendimentos não oriundos de trabalho, por condição de imputação, segundo as variáveis indicadoras de imputação dos rendimentos

Variáveis indicadoras de imputação dos rendimentos	Pessoas registradas nas variáveis de rendimentos não oriundos de trabalho		
	Total	Condição de imputação	
		Sem imputação	Com imputação
	Valores absolutos		
M4573 - Aposentadoria e pensão	16 194 309	16 081 787	112 522
M4583 - Aluguel	16 194 309	16 099 246	95 063
M4593 - Pensão alimentícia, mesada ou doação de não morador	16 194 309	16 097 014	97 295
M4603 - Renda mínima, bolsa-escola (programas oficias de auxílio)	16 194 309	16 089 242	105 067
M4613 - Outros rendimentos	16 194 309	16 080 151	114 158
	Valores relativos		
M4573 - Aposentadoria e pensão	100,0	99,3	0,7
M4583 - Aluguel	100,0	99,4	0,6
M4593 - Pensão alimentícia, mesada ou doação de não morador	100,0	99,4	0,6
M4603 - Renda mínima, bolsa-escola (programas oficias de auxílio)	100,0	99,4	0,6
M4613 - Outros rendimentos	100,0	99,3	0,7

Fonte: IBGE, Censo Demográfico 2000.

Ocupação e Atividade Econômica

No questionário do Censo Demográfico 2000, as informações sobre ocupação e atividade econômica são descritas pelo informante e registradas por extenso. Na apuração, estas descrições foram associadas a categorias definidas nas classificações de ocupações e de atividades econômicas, por meio de códigos numéricos. A imputação de informações nessas variáveis é posterior ao processo de codificação.

As críticas referentes à parte de trabalho e rendimento estão divididas em dois aplicativos do sistema DIA. Concluída a execução do primeiro aplicativo, as variáveis nele tratadas não podem mais ser alteradas no segundo aplicativo.

As variáveis V0447 “posição na ocupação” e V4452 “ocupação” foram tratadas no primeiro aplicativo, enquanto que a variável V4462 (atividade econômica) fez parte do segundo aplicativo. Assim sendo, no caso de ter havido inconsistência entre o registro da variável V4462 e os das variáveis V0447 e V4452, o procedimento adotado foi o de ajustar o registro da variável V4462.

Tabela 12.45 - Pessoas registradas nas variáveis V4452 e V4462, por condição de imputação, segundo as variáveis

Variáveis	Pessoas registradas nas variáveis V4452 e V4462		
	Total	Condição de imputação	
		Sem imputação	Com imputação
	Valores absolutos		
V4452 - ocupação	7 782 899	7 446 237	336 332
V4462 - atividade	7 782 899	6 746 051	1 036 848
	Valores relativos		
V4452 - ocupação	100,0	95,7	4,3
V4462 - atividade	100,0	86,7	13,3

Fonte: IBGE, Censo Demográfico 2000.

A proporção de informações imputadas na variável V4452 “ocupação” foi de 4,3%. Desse grupo, 14,6% (ou 49 081 pessoas) foram classificadas no processo de codificação como tendo ocupação "ignorada" e passaram a ter a informação de que sua ocupação estava mal definida ou era ignorada (código "zero"). Outras 7,7% (25 959 pessoas) tinham originalmente código de ocupação em branco. As demais pessoas sofreram imputação no código de ocupação como resultado do processo de crítica.

Já os registros da variável de atividade econômica (V4462) tiveram maior proporção de informações imputadas (13,3% dos códigos ou 1 036 848 pessoas). Desses, 8,5% (87 997 pessoas) estavam originalmente em branco. Um grupo majoritário de 76,3% (ou 791 307 pessoas) não tinha marca de imputação nem na variável V0447 "posição na ocupação", nem na variável V4452 "ocupação", ou seja, possuía alguma resposta nessas questões. Outros 17,2% sofreram imputação no código de ocupação, mas declararam sua posição na ocupação.

Tabela 12.46 - Pessoas registradas com imputação na variável V4462, por condição de imputação na variável V4452, segundo a condição de imputação na variável V0447

Condição de imputação na variável V0447	Pessoas registradas com imputação na variável V4462		
	Total	Condição de imputação	
		Sem imputação	Com imputação
	Valores absolutos		
Total	**1 036 848**	**849 427**	**187 421**
Sem imputação	969 692	791 307	178 385
Com imputação	67 156	58 120	9 036
	Valores relativos		
Total	**100,0**	**81,9**	**18,1**
Sem imputação	93,5	76,3	17,2
Com imputação	6,5	5,6	0,9

Fonte: IBGE, Censo Demográfico 2000.

Com a finalidade de explicar essa imputação de 13,3% na variável atividade econômica, foi construída a matriz de contingência para o Brasil. Nela, foram identificados os maiores fluxos de imputação entre grupos de atividade econômica. Em uma análise integrada por especialistas de todas as áreas envolvidas, considerou-se razoável a imputação feita através do sistema DIA, tendo em vista que as atividades imputadas passaram a ter coerência com a ocupação e posição na ocupação declaradas pelo informante. Declarações estas que também foram confirmadas, quando comparadas com as imagens digitalizadas dos questionários. O mesmo procedimento foi utilizado para confirmar as imputações de ocupação.

Outras variáveis relacionadas com os temas "Trabalho e rendimento"

Nas demais variáveis do grupo temático "Trabalho e rendimento", as proporções de informações imputadas ficaram, majoritariamente, em torno de 1%.

Os valores absolutos e as proporções de informações imputadas nas demais variáveis do tema estão na Tabela 12.46 que segue.

Tabela 12.47 - Pessoas registradas em outras variáveis relacionadas com o tema "Trabalho e rendimento", por condição de imputação, segundo as variáveis

Variáveis	Pessoas registradas nas variáveis V4452 e V4462		
	Total	Condição de imputação	
		Sem imputação	Com imputação
Valores absolutos			
V0439 - Trabalho remunerado na semana	16 194 309	16 056 004	138 305
V0440 - Estava afastado temporariamente do trabalho	9 463 867	9 192 840	271 027
V0441 - Trabalho não remunerado em ajuda a conta própria ou empregador	9 143 596	9 034 042	109 554
V0442 - Trabalho não remunerado em ajuda empregado	8 880 062	8 727 384	152 678
V0443 - Trabalhador produção próprio consumo	8 721 092	8 613 347	107 745
V0448 - Empregados pelo Regime Jurídico dos Funcionários Públicos ou como Militares	1 930 353	1 890 701	39 652
V0449 - Quantidade de empregados	209 478	207 857	1 621
V0450 - Constribuição para previdência	3 980 423	3 945 532	34 891
V0453 - Horas trab. principal na semana	7 782 899	7 718 116	64 783
V0455 - Providência para conseguir trabalho	8 411 410	8 333 781	77 629
V0456 - Aposentado de inst. previdência oficial	16 194 309	15 896 974	297 335
Valores relativos			
V0439 - Trabalho remunerado na semana	100,0	99,1	0,9
V0440 - Estava afastado temporariamente do trabalho	100,0	97,1	2,9
V0441 - Trabalho não remunerado em ajuda a conta própria ou empregador	100,0	98,8	1,2
V0442 - Trabalho não remunerado em ajuda a empregado	100,0	98,3	1,7
V0443 - Trabalhador produção próprio consumo	100,0	98,8	1,2
V0448 - Empregados pelo Regime Jurídico dos Funcionários Públicos ou como Militares	100,0	97,9	2,1
V0449 - Quantidade de empregados	100,0	99,2	0,8
V0450 - Constribuição para previdência	100,0	99,1	0,9
V0453 - Horas trab. principal na semana	100,0	99,2	0,8
V0455 - Providência para conseguir trabalho	100,0	99,1	0,9
V0456 - Aposentado de inst. previdência oficial	100,0	98,2	1,8

Fonte: IBGE, Censo Demográfico 2000.

Fecundidade e Mortalidade Infantil

As informações referentes à Fecundidade e Mortalidade Infantil constituem um único bloco no questionário do Censo Demográfico 2000, portanto, os procedimentos para a análise das variáveis relacionadas a esses temas foram semelhantes.

Em um primeiro momento, foram descobertos erros sistemáticos no campo, tais como informações deixadas em branco e valores fora dos limites, e ensaiaram-se diversos procedimentos de imputação determinística para corrigi-los. Detectaram-se, também, problemas na leitura realizada pelo *scanner* que afetavam os níveis de mortalidade infantil, corrigidos mediante a aplicação dos filtros descritos no anexo de CD-ROM deste texto.

Quando do cálculo das taxas de mortalidade infantil, utilizando informações preliminares da amostra do Censo 2000 sobre filhos tidos nascidos vivos e filhos sobreviventes, ocorreram divergências ao comparar os resultados das estimativas obtidas com as esperadas. Essa comparação tomou como parâmetros as estimativas derivadas do Projeto de Projeções de População por sexo e idade para o Brasil, Grandes Regiões e Unidades da Federação, do IBGE, realizado em convênio com o Fundo da População das Nações Unidas – FNUAP, que utilizou a metodologia dos Métodos das Componentes e as decorrentes da incorporação das séries da Pesquisa Nacional por Amostra de Domicílios – PNAD – da década de 1990, à série histórica obtida via Censos Demográficos 1940 a 1991.

Em decorrência da extrema sensibilidade das estimativas de mortalidade às variações nos dados originais e das diferenças observadas nos dados básicos sem crítica, criou-se dentro do âmbito da Diretoria de Pesquisas um grupo especial de trabalho para analisar o tratamento das informações sobre fecundidade das mulheres e mortalidade de seus respectivos filhos. Esse grupo foi integrado por demógrafos deste Instituto, por estatísticos e especialistas na área de metodologia. As primeiras estimativas já apontavam para uma diferença nas taxas de mortalidade infantil, que apresentava aumento em relação às séries observadas a partir das PNADs realizadas na década de 1990. Visando aprofundar a investigação desse problema, foi solicitado ao Departamento de Metodologia um estudo sobre a confiabilidade e a variabilidade das estimativas da mortalidade infantil, baseado nos dados da PNAD.

Esse estudo (SILVA; PESSOA, 2002) esclareceu o porquê das diferenças observadas entre os dados do Censo e da PNAD. Em relação a esta última, o fato de a amostra ser sempre escolhida a partir dos mesmos setores e a estimativa analisada ser muito sensível a pequenas variações nos dados básicos, causava diferenças nos valores obtidos. Embora essa tendência seja correta para essa amostra, não se ajusta completamente ao comportamento da mortalidade no Brasil.

Com base nas conclusões do grupo de trabalho, decidiu-se que as informações de mortalidade e fecundidade seriam divulgadas em volume separado, com os cálculos dos indicadores efetuado a partir da amostra completa do Censo Demográfico 2000.

Com respeito ao tratamento da informação para a divulgação definitiva dos resultados, os exercícios realizados apontaram para uma relativa aproximação entre os valores das taxas e os valores esperados, porém mostraram de forma igual que o processo de crítica e imputação poderia não ser suficiente para corrigir os problemas detectados, especialmente considerando que as correções se restringiam aos dados que apresentavam inconsistências entre si, isto é, que as informações mencionadas já estavam dentro do intervalo definido para cada variável ou conjunto de variáveis. Nesse sentido, os filtros e procedimentos de verificação solicitados para a etapa da entrada de dados foram efetivamente aplicados

antes que o conjunto das informações fosse submetido aos procedimentos de crítica através do DIA. Posteriormente, aplicaram-se as imputações determinísticas para resolver os problemas mais freqüentes na informação básica.

Ao final do processo de crítica e imputação, concluiu-se sobre as estimativas da Fecundidade e da Mortalidade Infantil que: as distribuições das variáveis que permitiam estimar os citados parâmetros demográficos não sofreram alterações com o processo de crítica e imputação; e as estimativas propriamente ditas não foram modificadas substancialmente com a passagem do sistema DIA.

Resumindo, na divulgação dos resultados definitivos, foram utilizados procedimentos normais de crítica e imputação já com os dados depurados, chegando-se a resultados finais não muito distintos daqueles divulgados inicialmente, na ocasião do lançamento da *Tabulação avançada do censo demográfico 2000: resultados preliminares da amostra (2002).*

Ainda analisando os dados do bloco de Fecundidade, observou-se que os níveis da natimortalidade, expressos pelos percentuais de "filhos tidos nascidos mortos no total de filhos tidos", permaneceram elevados ao longo de todo o processo de crítica e imputação. Tal fenômeno foi observado no Censo Demográfico 1980 e foi constatado que o mesmo não fora provocado pelo processo de imputação.

Relativamente às razões de sexo dos "filhos tidos nascidos vivos e nascidos mortos", segundo os grupos de idade das mulheres, os Gráficos 12.16 e 12.17 mostram de que modo a imputação através do sistema DIA, com as conseqüentes correções das inconsistências e imputação dos dados faltantes, modificou o comportamento desses indicadores.

No caso dos "filhos tidos nascidos vivos" houve, de modo geral, pequenas reduções nos valores correspondentes às razões de sexo, quando comparados os dados criticados com os não criticados. Estes valores oscilaram em torno de 105 homens para cada 100 mulheres, o que se aproxima do padrão observado nas Estatísticas Vitais. A série das razões de sexo no nascimento de 2000 mostrou-se semelhante à de 1980, tendo 1991 um comportamento atípico. Entretanto, na variável "filhos tidos nascidos mortos", as razões finais ficaram bastante próximas das iniciais, oscilando próximo a 150 homens para cada 100 mulheres. A exceção foi do grupo de 10 a 14 anos de idade que, por ser um grupo muito rarefeito, apresentou grande variação. Na comparação com 1980 e 1991, observou-se novamente uma semelhança com os dados do Censo Demográfico 1980, sendo que para 1991, essas razões de sexo eram mais elevadas que aquelas apresentadas internacionalmente sobre natimortalidade.

Gráfico 12.16 - Razões de sexo dos filhos tidos nascidos vivos Brasil - 1980-2000

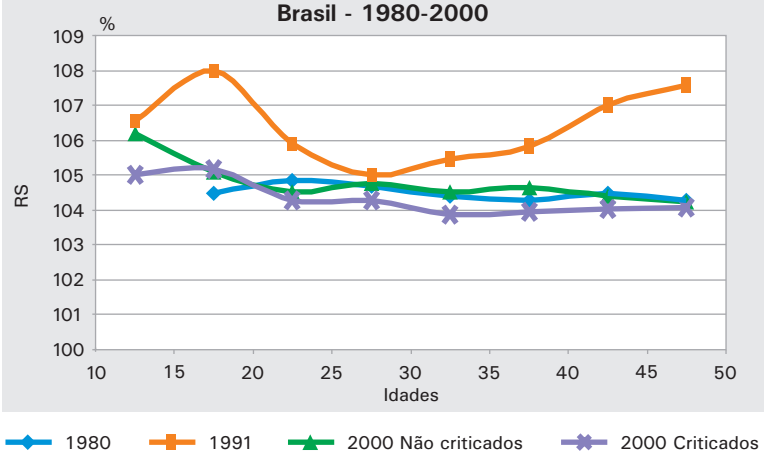
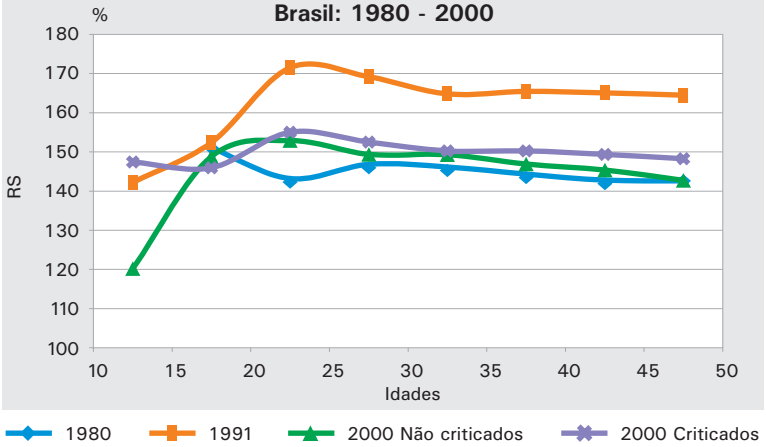


Gráfico 12.17 - Razões de sexo dos filhos tidos nascidos mortos
Brasil: 1980 - 2000



12.5.2.4 Imputação das variáveis de rendimento

Este item descreve o processo de imputação das variáveis de rendimentos dos indivíduos recenseados no Censo Demográfico 2000 através do Questionário da Amostra. As motivações que levaram ao desenvolvimento de tal processo e uma breve descrição da metodologia já estão apresentadas no item 12.3.2.7, que trata da imputação da variável de rendimento do responsável pelo domicílio, ou individual em domicílio coletivo, do Conjunto Universo. Aqui, são apresentados os aspectos específicos dessa aplicação e os resultados obtidos.

Vale lembrar que o Questionário da Amostra foi aplicado a 20% ou 10% dos domicílios de cada um dos municípios brasileiros, respectivamente, com até 15 000 e com mais de 15 000 habitantes. Nesse questionário, havia uma série de perguntas cujo objetivo era caracterizar os moradores do domicílio em 31 de Julho de 2000, data de referência do Censo Demográfico 2000. Entre essas perguntas, havia um grupo referente aos valores e às origens dos rendimentos dos moradores com 10 anos ou mais de idade, na data de referência do censo. Desse grupo de perguntas, derivou-se um conjunto de variáveis que descrevem o perfil dos rendimentos, no mês de julho de 2000, de cada um dos moradores recenseados. Essas variáveis são as seguintes:

Rendimento bruto no trabalho principal;

Rendimento bruto nos demais trabalhos;

Rendimentos provenientes de aposentadoria, pensão;

Rendimentos provenientes de aluguel;

Rendimentos provenientes de pensão alimentícia, mesada, doação recebida de não-morador;

Rendimentos provenientes de renda mínima/bolsa-escola, seguro-desemprego, etc. (programas oficiais de auxílio);

Outros rendimentos recebidos; e

Rendimento total (soma de todos os rendimentos obtidos).

A existência de valores faltantes nessas variáveis (valores não declarados) pode trazer diversos efeitos sobre análises de rendimentos a serem feitas. Daí, fez-se necessária a imputação de valores de rendimentos dos não respondentes.

Aspectos gerais

Algumas colocações a respeito dos efeitos da não-resposta em pesquisas já foram tratadas no item 12.3.2.7. Uma questão ali colocada refere-se aos possíveis efeitos de ignorar a não resposta ao se fazer inferências sobre uma variável estudada. Como resposta a essa questão, é visto que, em caso de não-resposta não diferencial, isto é, ao acaso, o efeito existente é o de aumento da variância das estimativas. Já para o caso da não-resposta diferencial, há o efeito de um impacto que se dá sob a forma de vício nas estimativas obtidas, com esse vício crescendo com a taxa de não-resposta e com a diferença entre respondentes e não respondentes.

Análises realizadas com dados do Censo Demográfico 1991 mostraram que a não-resposta nos rendimentos dos chefes de domicílio ocorria de forma diferencial em relação a algumas das variáveis presentes no questionário, brevemente descritas em 12.3.2.7.

Estudo semelhante foi feito para o projeto de Imputação de Rendimentos no Questionário da Amostra, onde a partir dos microdados utilizados para a Tabulação Avançada do Censo Demográfico 2000, verificou-se a existência de não-resposta diferencial nos quesitos de rendimento. Um exemplo disso é apresentado na Tabela 12.48, onde são observadas taxas de não-resposta do rendimento total em algumas unidades da federação, segundo algumas categorias da variável relação com o responsável pelo domicílio. Conforme pode-se verificar nessa tabela, no contigente de não respondentes do rendimento total (moradores que não responderam a uma ou mais das categorias de rendimento), há uma alta concentração na categoria "filho(a), enteado(a)", caracterizando-se assim a não resposta diferencial para o rendimento total com respeito à relação com o responsável pelo domicílio.

Tabela 12.48 – Distribuição da não resposta da variável de rendimento total, para as pessoas de 10 anos ou mais de idade, em algumas das categorias da variável relação de parentesco com o responsável pelo domicílio, para algumas Unidades da Federação

Unidades da Federação	Distribuição da não resposta da variavel rendimento total, por categoria da variável relação com o responsável pelo domicílio (%)				
	Todas as categorias	Pessoa responsável	Cônjuge, Companheiro(a)	Filho(a), Enteado(a)	Demais categorias
Pará	100	4	2	79	15
Tocantins	100	16	16	60	8
Bahia	100	1	9	74	16
São Paulo	100	5	8	75	12
Rio Grande do Sul	100	0	8	75	17
Distrito Federal	100	0	0	75	25

Fonte: IBGE, Censo Demográfico 2000, Tabulação Avançada.

São duas as alternativas para lidar com o problema da não resposta diferencial: uso de estimadores adequados para dados faltantes (LITTLE; RUBIN, 1987) e de métodos de imputação (substituição por valores estimados em cada

caso individual). No caso de não-resposta parcial de uma variável, a preferência das agências de estatísticas oficiais é geralmente por métodos baseados em imputação das variáveis não informadas. Tal preferência se deve à maior simplicidade dessa alternativa no processamento posterior dos dados, particularmente quando estes precisam ser publicados na forma de arquivos de microdados com as informações de cada respondente individual. Albieri (1992) investigou a aplicação de vários métodos para imputação da renda na Pesquisa Mensal de Emprego do IBGE.

O método desenvolvido para a imputação de rendimentos no Questionário da Amostra trabalha com a idéia de estabelecer uma relação entre rendimentos declarados pelos moradores e um grupo de variáveis do questionário da amostra, cujos valores são conhecidos para todos os moradores dos domicílios pesquisados, e, a partir dessa relação, imputar valores de rendimento para os não respondentes.

As variáveis existentes no Questionário da Amostra foram estudadas para a seleção de variáveis explicativas dos rendimentos dos moradores, em suas categorias e em seu total, buscando-se um conjunto que fosse diverso o suficiente para descrever de forma satisfatória as diferentes relações existentes ao longo do País. Infelizmente, verificou-se que relações satisfatórias entre rendimentos e variáveis do Questionário da Amostra só eram encontradas para o rendimento do trabalho principal e o rendimento total, isto é, não foram verificadas relações entre as variáveis do questionário e as demais categorias de rendimento. Com isso, para a imputação das categorias de rendimento, que não a do trabalho principal e a de total, foi adotado um procedimento baseado na relação entre o rendimento total e as variáveis selecionadas do Questionário da Amostra.

Ao contrário do ocorrido na imputação de renda dos responsáveis por domicílios ou individuais em domicílios coletivos no Conjunto Universo do Censo Demográfico 2000, na imputação de rendimentos no Questionário da Amostra não foi realizada a imputação de rendimentos nulos. Isto porque há variáveis no Questionário da Amostra que poderiam ser utilizadas para fins de predição de quem teria ou não rendimento nulo nas categorias de rendimento existentes, o que não ocorria no caso do Conjunto Universo.

A seguir, é apresentado o conjunto de variáveis selecionadas para utilização no processo de imputação, com a descrição de cada variável precedida da respectiva nomenclatura adotada:

1. COND.TRAB – associada à condição do morador no seu trabalho principal na semana de 23 a 29 de julho de 2000, possuindo as seguintes categorias: a – trabalhador doméstico com carteira de trabalho assinada; b – trabalhador doméstico sem carteira de trabalho assinada; c – empregado com carteira de trabalho assinada; d – empregado sem carteira de trabalho assinada, que não militar ou funcionário público estatutário; e – militar ou funcionário público estatutário; f – empregador; g – conta-própria; h – aprendiz ou estagiário sem remuneração; i – não remunerado em ajuda a membro do domicílio; j – trabalhador na produção para o próprio consumo;

2. GRUP.ATIV – resultante de agregação das categorias referentes à atividade principal do negócio, firma, empresa, instituição ou entidade em que o morador trabalhava na semana de 23 a 29 de julho de 2000;

3. TOT.BAN – resultante da uma combinação entre a quantidade de banheiros existentes no domicílio e a existência ou não de sanitário em domicílios sem banheiro. Assume os seguintes valores: -1, em domicílios sem banheiro e sanitário; 0, em domicílios sem banheiro e com sanitário; total de banheiros, em domicílios com a existência de 1 ou mais banheiros;

4. UTENS – resultante da existência ou não de determinados utensílios domésticos no domicílio, assumindo os seguintes valores: 0, em domicílios sem nenhum dos seguintes utensílios: videocassete, máquina de lavar roupa, forno de microondas e computador; 1, em domicílios com pelo menos um dos utensílios citados;

5. QTD.TVS – associada à quantidade de televisores existentes no domicílio, assumindo os seguintes valores: 0, em domicílios sem televisor; 1, em domicílios com 1 televisor; 2, em domicílios com 2 televisores; 3, em domicílios com 3 ou mais televisores;

6. QTD.AUTO – associada à quantidade de automóveis para uso particular existentes no domicílio, assumindo os seguintes valores: 0, em domicílios sem automóvel para uso particular; 1, em domicílios com 1 automóvel para uso particular; 2, em domicílios com 2 automóveis para uso particular; 3, em domicílios com 3 ou mais automóveis para uso particular;

7. QTD.AR – associada à quantidade de aparelhos de ar-condicionado, assumindo os seguintes valores: 0, em domicílios sem aparelho de ar-condicionado; 1, em domicílios com 1 aparelho de ar-condicionado; 2, em domicílios com 2 aparelhos de ar-condicionado; 3, em domicílios com 3 ou mais aparelhos de ar-condicionado;

8. SEXO – associada ao sexo do morador, possuindo as seguintes categorias: a - morador do sexo masculino; b - morador do sexo feminino;

9. REL.RESP.DOM – associada à relação do morador com o responsável pelo domicílio, possuindo as seguintes categorias: a – pessoa responsável; b – cônjuge, companheiro(a); c – filho(a), enteado(a); d – pai, mãe, sogro(a); e – neto(a), bisneto(a); f – irmão, irmã; g – outro parente; h – agregado(a); i – pensionista; j – empregado(a) doméstico(a); k – parente do(a) empregado(a) doméstico(a); l – individual em domicílio coletivo;

10. IDADE – idade do morador, em anos completos, em 31 de julho de 2000;

11. ANOS.EST – anos de estudo do morador;

12. IND.TRAB.PRINC – indicativa da condição do morador em relação ao rendimento proveniente de trabalho principal, possuindo as seguintes categorias: a – o rendimento proveniente de trabalho principal é zero; b – somente possui benefícios; c – possui rendimento proveniente de trabalho principal; d – não possui trabalho principal ou trabalha na produção para o próprio consumo;

13. IND.APOSENT – indicativa da condição do morador em relação ao rendimento proveniente de aposentadoria/pensão, possuindo as seguintes categorias: a – não possui rendimento proveniente de aposentadoria/pensão; b – possui rendimento proveniente de aposentadoria/pensão.

Metodologia

Conforme dito anteriormente, apenas para o rendimento total e para o rendimento do trabalho principal foi possível obter outras variáveis - covariáveis - que tivessem poder de descrição do comportamento dessas variáveis. Para as demais categorias de rendimento, o procedimento de imputação foi realizado com a utilização da árvore de regressão definida com a variável rendimento total. Assim, cada indivíduo era localizado em um estrato da árvore do rendimento total e, nesse estrato, selecionava-se o doador do rendimento não declarado. Em caso de não declaração de mais de um rendimento, as informações não eram selecionados de um único doador, pois havia o risco, considerável, deste não ser encontrado com tais rendimentos não nulos. Em caso de não haver doador para alguma categoria de rendimento em algum estrato, o doador era selecionado no subgrupo de registros que originou o estrato.

Aplicação e conclusões

Para o processo de imputação de rendimentos, utilizaram-se os mesmos lotes de registros definidos para a crítica e imputação de dados nos questionários de domicílios selecionados na amostra do Censo Demográfico 2000, realizadas com o uso do sistema DIA. Esses lotes correspondem a uma partição do conjunto de moradores dos domicílios selecionados na amostra, obedecendo aos domínios das unidades da federação (UF), isto é, um mesmo lote não contém registros de mais de uma UF. Para o processo de imputação de rendimentos, foram excluídos de cada lote os registros cujo rendimento total encontrava-se fora das cercas construídas para detectar outliers (valores atípicos). Nos 215 lotes utilizados na imputação de rendimentos, havia o total de 16.130.468 registros, com o menor deles possuindo 4.696 registros e o maior 295.074 registros. A distribuição da quantidade de lotes por UF pode ser vista na Tabela 12.16, deste capítulo.

Tal como no procedimento de imputação da variável de rendimento apurada no Conjunto Universo, para cada um dos 215 lotes de apuração dos questionários da amostra foi aplicado o procedimento de imputação baseado em árvores de regressão, descrito anteriormente. Tal como no procedimento de imputação da variável de rendimento do Conjunto Universo descrito no item 12.3, esse procedimento foi implementado com o software S-Plus e executado em ambiente operacional Windows 98. Como os lotes de registros residiam em arquivos do ambiente operacional OS/390 (mainframe IBM), foi desenvolvida, utilizando o software SAS e seus recursos para a conexão desses dois ambientes operacionais, uma rotina computacional para automatizar todo o processo de produção dessa imputação, constituído das seguintes etapas: a) preparação do arquivo de entrada para o S-Plus; b) ativação do S-Plus para a imputação propriamente dita; c) transferência dos resultados para o ambiente OS/390; e d) atualização dos registros nos lotes originais com os valores imputados.

A regra de parada na construção das árvores de regressão baseou-se no número máximo de nós terminais permitido nas árvores e no contingente populacional mínimo exigido em cada nó terminal. Visto que seria impraticável a análise dos gráficos de queda da deviance para cada uma das 215 árvores, tal como descrito no item 12.3.2.7, uma das regras de parada adotada foi a da partição de cada lote em no máximo 25 estratos. Análises preliminares com dados da Tabulação Avançada do Censo 2000, anteriormente citadas, indicaram ser este um número de nós terminais para o qual, em geral, não haveria "ganhos consideráveis" com novas partições. Quanto ao contingente populacional em cada estrato, foi estipulado que este deveria ser de no mínimo 100 pessoas.

Estatísticas referentes às taxas de não-resposta nos lotes para cada uma das categorias de rendimento e para o rendimento total são apresentadas nas Tabelas 12.49 a 12.56. As taxas de não-resposta para cada categoria de rendimento foram calculadas, levando-se em conta apenas os moradores que possuíam cada um dos rendimentos; somente para o rendimento total, foram considerados todos os moradores presentes na amostra.

Tabela 12.49 - Estatísticas descritivas das taxas de não-resposta nos lotes de imputação para rendimento no trabalho principal

Taxas de não-resposta - rendimento no trabalho principal					
Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo
0,68%	1,91%	3,05%	4,15%	5,42%	18,01%

Fonte: IBGE, Censo Demográfico 2000.

Tabela 12.50 - Estatísticas descritivas das daxas de não-resposta nos lotes de imputação para rendimento nos demais trabalhos

Taxas de não-resposta - rendimento nos demais trabalhos					
Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo
0%	%	%	1,65%	0,16%	5,77%

Fonte: IBGE, Censo Demográfico 2000.

Tabela 12.51 - Estatísticas descritivas das taxas de não-resposta nos lotes de imputação para rendimento proveniente de aposentadoria, pensão

Taxas de não-resposta - rendimento proveniente de aposentadoria, pensão					
Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo
0,87%	1,74%	2,33%	2,49%	2,96%	7,03%

Fonte: IBGE, Censo Demográfico 2000.

Tabela 12.52 - Estatísticas descritivas das taxas de não-resposta nos lotes de imputação para rendimento proveniente de aluguel

Taxas de não-resposta - rendimento proveniente de aluguel					
Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo
0%	0,16%	0,43%	0,48%	0,70%	2,81%

Fonte: IBGE, Censo Demográfico 2000.

Tabela 12.53 - Estatísticas descritivas das taxas de não-resposta nos lotes de imputação para rendimento proveniente de pensão alimentícia, mesada, doação recebida de não-morador

Taxas de não-resposta - rendimento proveniente de pensão alimentícia, mesada, doação recebida de não-morador					
Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo
0%	0,11%	0,38%	0,45%	0,67%	2,77%

Fonte: IBGE, Censo Demográfico 2000.

Tabela 12.54 - Estatísticas descritivas das taxas de não-resposta nos lotes de imputação para rendimento proveniente de programas oficiais de auxílio

Taxas de não-resposta - rendimento proveniente de programas oficiais de auxílio					
Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo
0%	1,66%	3,03%	4,69%	5,63%	41,68%

Fonte: IBGE, Censo Demográfico 2000.

Tabela 12.55 - Estatísticas descritivas das taxas de não-resposta nos lotes de imputação para outros rendimentos

Taxas de não-resposta – outros rendimentos					
Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo
0%	0,31%	0,57%	0,62%	0,82%	2,56%

Fonte: IBGE, Censo Demográfico 2000.

Tabela 12.56 - Estatísticas descritivas das taxas de não-resposta nos lotes de imputação para rendimento total

Taxas de não-resposta - rendimento total					
Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo
0,59%	1,25%	1,63%	1,93%	2,43%	5,65%

Fonte: IBGE, Censo Demográfico 2000.

Os gráficos 12.18 e 12.19 apresentam os resultados da participação de cada variável explicativa nas árvores de regressão construídas para o rendimento total e para o rendimento do trabalho principal, respectivamente. Essa participação é vista sob o ponto de vista da ordem em que a variável gerou uma partição na árvore pela primeira vez. As variáveis são identificadas nos gráficos de acordo com a numeração atribuída, quando foram apresentadas no início deste item.

Gráfico 12.18 - Posições de entrada das variáveis auxiliares nas Árvores de Rendimento Total

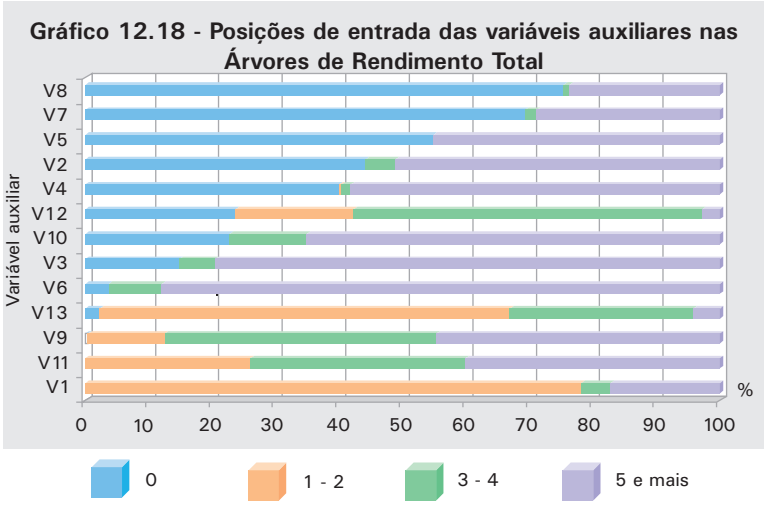
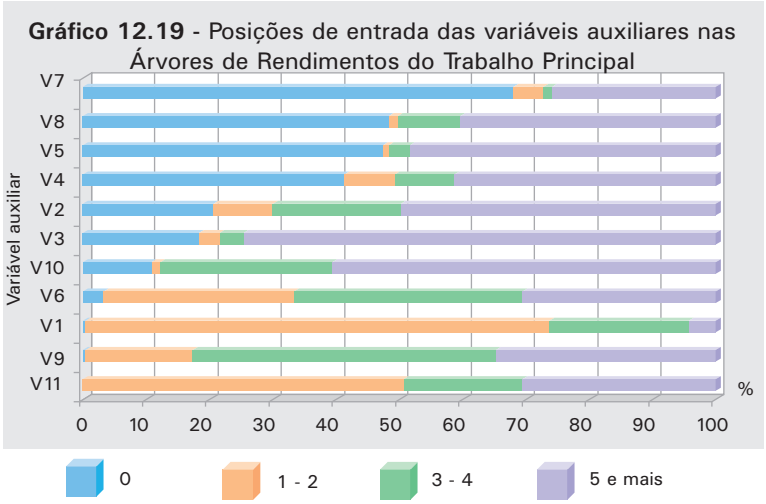


Gráfico 12.19 - Posições de entrada das variáveis auxiliares nas Árvores de Rendimentos do Trabalho Principal



Como forma de avaliar a qualidade do resultado da imputação em cada lote, foi aplicado o teste estatístico de Kolmogorov-Smirnov (LEHMANN, c1975). Este teste visa verificar se duas amostras de dados provém de uma mesma população. O teste foi aplicado para se comparar os vetores de rendimentos totais em cada subgrupo antes e depois da execução do procedimento de imputação. Cada lote só teve seu respectivo processo de imputação aprovado se o teste de Kolmogorov-Smirnov indicasse que as rendas antes e depois da imputação, em cada estrato formado, apresentavam a "mesma distribuição".

Finalizando, são apresentadas na Tabela 12.56 algumas estatísticas referentes às distribuições das taxas de imputação nos estratos obtidos nas árvores de regressão construídas para o rendimento total.

Tabela 12.57 - Estatísticas descritivas das taxas de não-resposta do rendimento total nos estratos das árvores de regressão dos lotes de imputação de rendimento no questionário da amostra

Taxas de não resposta nos estratos das árvores de regressão do rendimento total					
Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo
0%	1,25%	2,48%	4,72%	4,60%	100%

Fonte: IBGE, Censo Demográfico 2000.

12.5.3 Expansão da amostra

Para expansão dos dados coletados no Questionários da Amostra do Censo Demográfico 2000, foram calculados pesos para as unidades domiciliares pesquisadas, sendo tais pesos atribuídos também a cada um de seus moradores. Por unidades domiciliares pesquisadas, entende-se os domicílios particulares ocupados e as famílias e pessoas sós moradoras em domicílio coletivo.

12.5.3.1 Método para obtenção dos Pesos

O método utilizado para obtenção dos pesos foi um processo de calibração em relação a um conjunto de variáveis auxiliares (restrições) para as quais se conhecem os totais populacionais, já que tais variáveis auxiliares foram levantadas pelo Questionário Básico. A calibração buscou ajustar os pesos iniciais (inverso da

fração amostral de domicílios) de maneira que, dentro de uma determinada área geográfica, denominada área de ponderação, ao se aplicar os pesos calibrados às variáveis auxiliares, fossem obtidos os totais já conhecidos para todas as unidades da população que constituem o universo da pesquisa. Dessa maneira, além de melhorar a precisão dos estimadores, obtêm-se estimativas mais consistentes para as variáveis pesquisadas somente pelo questionário da amostra.

O cálculo dos pesos calibrados foi baseado no método dos Mínimos Quadrados Generalizados – MQG, porém com a imposição de limites nos pesos finais, para evitar pesos muito pequenos ou muito grandes. O limite mínimo utilizado foi 1, de maneira que um domicílio representasse pelo menos ele próprio. O limite máximo foi definido como cinco vezes o peso médio esperado, ou seja, 25 no caso de municípios com fração amostral planejada de 20% (caso em que o peso médio esperado era 5) e 50 no caso de municípios com fração amostral planejada de 10% (caso em que o peso médio esperado era 10). Sem a utilização desses limites, o método MQG pode gerar pesos negativos ou muito grandes, o que não teria sentido prático.

A metodologia para utilização do método MQG baseou-se em proposta de Bankier (1990) e, para sua implementação, um sistema em linguagem SAS foi desenvolvido por técnicos do IBGE.

O produto final da aplicação dessa metodologia é um peso ajustado para cada unidade domiciliar da amostra, ou seja, cada um dos questionários da amostra, que é repetido nos registros de cada pessoa moradora na unidade domiciliar.

a) Definição das variáveis auxiliares

A escolha das variáveis auxiliares cujos valores são utilizados como restrições no processo de ajustamento do qual decorrem os pesos é um aspecto importante do método aplicado. A forma ou prioridade de tratamento dessas variáveis, sobretudo, quando não existe uma solução que atenda simultaneamente a todas as restrições, é outro ponto sensível do método.

As variáveis auxiliares constituem um subconjunto das variáveis comuns à amostra e ao universo e são referentes a características de domicílios ou de pessoas, apesar do ajustamento ser realizado de forma a fornecer pesos para cada uma das unidades domiciliares.

A metodologia de ajuste de um modelo linear generalizado multivariado envolve cálculos com matrizes, inclusive inversão. Por essa razão, as restrições definidas que, por sua vez, dão origem a uma dessas matrizes, devem satisfazer algumas condições essenciais, sendo a principal delas a de não serem linearmente dependentes (redundantes). Além disso, é também considerado o conceito de restrições quase linearmente dependentes (e, portanto, quase redundantes), que afetam a estabilidade da solução do modelo.

Outras duas condições impostas para a aplicação dessa metodologia referem-se à sua significância estatística. O tamanho da restrição, medido como o número de domicílios aos quais a restrição se aplica em uma dada área de ponderação, não deve ser muito pequeno sob pena de tornar instável o processo de estimação. Quando uma restrição não atinge um número mínimo de unidades domiciliares, fixado em função da fração de amostragem, essa restrição é considerada rara.

Além disso, uma restrição definida pode causar a obtenção de um peso muito grande ou muito pequeno, quando comparado com o peso médio esperado (5 ou 10) em função da fração amostral adotada na área de ponderação ou até um peso negativo, constituindo-se em restrição geradora de peso extremo.

Dessa forma, o programa de ajuste do modelo incorpora procedimentos de eliminação de restrições que se enquadrem nas condições acima, observando a ordem que segue: restrições raras, restrições redundantes, restrições quase redundantes e restrições responsáveis por pesos extremos.

Convém ressaltar que a eliminação de restrições pode implicar diretamente no fato de não se ter a garantia da calibração desejada para as variáveis eliminadas para a presente área de ponderação.

As restrições inicialmente definidas para a aplicação da metodologia MQG, para cada uma das áreas de ponderação, encontram-se na relação abaixo. Elas constituem o conjunto denominado conjunto 1 de restrições para calibração.

Em unidades domiciliares (domicílios particulares ocupados + famílias ou pessoas sós em domicílios coletivos)

1. Número total de pessoas
2. Número total de unidades domiciliares
3. Número de pessoas do sexo masculino
4. Número de pessoas na faixa de idade de 0 a 4 anos
5. Número de pessoas na faixa de idade de 5 a 9 anos
6. Número de pessoas na faixa de idade de 10 a 14 anos
7. Número de pessoas na faixa de idade de 15 e 19 anos
8. Número de pessoas na faixa de idade de 20 a 24 anos
9. Número de pessoas na faixa de idade de 25 a 29 anos
10. Número de pessoas na faixa de idade de 30 a 34 anos
11. Número de pessoas na faixa de idade de 35 a 39 anos
12. Número de pessoas na faixa de idade de 40 a 44 anos
13. Número de pessoas na faixa de idade de 45 a 49 anos
14. Número de pessoas na faixa de idade de 50 a 59 anos
15. Número de pessoas na faixa de idade de 60 a 69 anos
16. Número de pessoas na faixa de idade de 70 anos ou mais
17. Número de pessoas do sexo masculino na faixa de idade de 0 a 4 anos
18. Número de pessoas do sexo masculino na faixa de idade de 5 a 9 anos
19. Número de pessoas do sexo masculino na faixa de idade de 10 a 14 anos
20. Número de pessoas do sexo masculino na faixa de idade de 15 a 19 anos
21. Número de pessoas do sexo masculino na faixa de idade de 20 a 24 anos

22. Número de pessoas do sexo masculino na faixa de idade de 25 a 29 anos

23. Número de pessoas do sexo masculino na faixa de idade de 30 a 34 anos

24. Número de pessoas do sexo masculino na faixa de idade de 35 a 39 anos

25. Número de pessoas do sexo masculino na faixa de idade de 40 a 44 anos

26. Número de pessoas do sexo masculino na faixa de idade de 45 a 49 anos

27. Número de pessoas do sexo masculino na faixa de idade de 50 a 59 anos

28. Número de pessoas do sexo masculino na faixa de idade de 60 anos ou mais

29. Número de pessoas moradoras na situação urbana

30. Número de pessoas do sexo feminino moradoras na situação urbana

31. Número de pessoas do sexo feminino moradoras na situação rural

Em domicílios particulares permanentes ocupados

32. Número de pessoas do sexo masculino que são chefes ou individuais

33. Número total de pessoas

34. Número total de domicílios

35. Número de domicílios urbanos

36. Número de domicílios com 1 ou 2 moradores

37. Número de domicílios com 3 moradores

38. Número de domicílios com 4 moradores

39. Número de domicílios com 5 moradores

40. Número de domicílios com 6 ou mais moradores

b) Análise da qualidade da calibração

As restrições, acima apresentadas, foram agrupadas em dez conjuntos alternativos que foram utilizados em ordem de prioridade. Esse procedimento foi adotado para garantir que alguma calibração fosse feita, mesmo que em um conjunto menor de características. O primeiro conjunto foi formado por todas as restrições, como listadas, e os demais formados pela agregação de faixas etárias, agregação de faixas de moradores por domicílio ou mesmo a retirada de grupos de restrições.

No cálculo dos pesos calibrados, para cada área de ponderação, foi utilizado inicialmente o conjunto de restrições número 1. Quando não se obteve uma solução satisfatória, a área foi processada novamente, utilizando o conjunto 2 e assim sucessivamente até o conjunto 10, caso anteriormente não tenha sido atingida a qualidade de ajuste adequada. A composição de cada um dos 9 conjuntos alternativos de restrições está apresentada em anexo no CD-ROM.

A análise da qualidade do ajuste (calibração) era feita automaticamente pelo sistema através das diferenças entre os valores populacionais conhecidos para as restrições e os valores estimados utilizando-se os pesos calculados. Para cada grupo de restrições, foram definidos limites específicos tolerados para essas diferenças.

Para as áreas de ponderação onde não ocorreu o ajuste para nenhum dos dez conjuntos de restrições, o sistema automaticamente escolheu o conjunto que proporcionou o melhor ajuste, no sentido de minimizar a soma dos quadrados das diferenças entre o valor conhecido das restrições e o valor estimado para essas mesmas restrições. Essa estatística foi calculada com base nas variáveis do conjunto 1 de restrições, para todos os dez conjuntos avaliados.

Deve-se ressaltar que para algumas áreas de ponderação onde houve um desequilíbrio forte entre a fração amostral de domicílios e a fração amostral de pessoas, ou seja, em áreas onde o número médio de pessoas por domicílio no universo e na amostra diferiram muito, pode ter ocorrido falta de ajuste na variável total de pessoas. Portanto, é possível a ocorrência, para alguns municípios, de divergência entre o valor do número de pessoas calculado através da expansão da amostra e o valor verificado na investigação do universo dos domicílios, que são os números oficiais do censo.

Do total de 9 336 áreas de ponderação definidas para o Brasil, mais de 91% teve solução para o conjunto 1 de restrições.

No final do processo, foi garantido que pelo menos a restrição "número total de domicílios" fosse respeitada para todas as áreas de ponderação. Nesse contexto, "número total de domicílios" iguala o número total de questionários e engloba os domicílios particulares ocupados mais as famílias e pessoas sós moradoras em domicílios coletivos.

12.5.3.2 Áreas de ponderação

Define-se Área de Ponderação como sendo uma unidade geográfica, formada por um agrupamento mutuamente exclusivo de setores censitários, para a aplicação dos procedimentos de calibração das estimativas com as informações conhecidas para a população como um todo.

Foram definidas, para todo o Brasil, 9 336 áreas de ponderação e, tal como nos censos anteriores, a metodologia de expansão da amostra foi aplicada independentemente para cada uma delas.

O tamanho dessas áreas, em termos de número de domicílios e de população, não pode ser muito reduzido, sob pena de perda de precisão de suas estimativas. As áreas de ponderação foram definidas considerando essa condição e, também, os níveis geográficos mais detalhados da base operacional, como forma de atender a demandas por informações em níveis geográficos menores que os municípios.

Os livros técnicos de amostragem definem procedimentos para a determinação de tamanhos de amostra considerando os requisitos de precisão estabelecidos para uma pesquisa. Para tanto, define-se a margem de erro aceitável para o estimador amostral, supondo que a amostra seria selecionada sob Amostragem Aleatória Simples (AAS). Considerando o objetivo de estimar uma média com um erro máximo relativo de kr ao nível de confiança de 95%, tem-se uma equação que relaciona o tamanho total da amostra desejada com os requisitos de precisão especificados.

A definição do tamanho das áreas de ponderação para o cálculo dos pesos de expansão da amostra do Censo demográfico 2000 foi feita considerando questões técnicas estreitamente relacionadas com as acima descritas. O tamanho mínimo definido para uma área de ponderação foi estabelecido em 400 domicílios particulares ocupados na amostra, por ser um valor aproximado ao encontrado nos cálculos de tamanho de amostra aleatória simples, quando se considera a intenção de estimar uma proporção (pequena) de 5%, com uma precisão relativa máxima fixada não muito exigente (40%) em uma população considerada grande, para os efeitos de aproximação nas fórmulas, e considerando um nível de confiança estabelecido em 95%, para a construção de intervalos de confiança.

A decisão de fixar o tamanho da amostra e não o tamanho da população da área de ponderação foi tomada com base no fato que a precisão de estimativas provenientes de pesquisas por amostragem está diretamente relacionada com o tamanho absoluto da amostra e não com a fração amostral (relação entre tamanho de amostra e tamanho da população). Assim, nos municípios onde foi decidido que seriam definidas áreas de ponderação em nível geográfico mais desagregado que o próprio município, foi considerada essa restrição de tamanho, com o objetivo de preservar a precisão de estimativas. Nos municípios onde foi considerada apenas uma área, o próprio município, a restrição de tamanho não pode ser aplicada, pois o tamanho da amostra foi uma decorrência da fração amostral definida antes da realização do censo. Nos municípios pequenos em que, em função da definição da fração amostral, o tamanho da amostra de domicílios resultou em valores menores que 400 unidades, é possível que um número significativo de estimativas tenha baixa precisão, medida em termos de erro amostral. Albieri (2003) apresenta mais considerações sobre essa definição e suas implicações.

Para o Censo 2000, foram usados métodos e sistemas automáticos de formação de áreas de ponderação que conjugam critérios tais como tamanho (para permitir estimativas com qualidade estatística em áreas pequenas), contigüidade (no sentido de serem constituídas por conjuntos de setores limítrofes com sentido geográfico) e homogeneidade em relação a um conjunto de características populacionais e de infra-estrutura conhecidas.

As áreas de ponderação foram criadas, considerando os seguintes critérios:

- maior nível geográfico utilizado é o município; isto significa que uma área de ponderação é composta por setores censitários dentro de um único município, podendo ser o próprio município;
- menor tamanho de uma área de ponderação não municipal é de 400 domicílios particulares ocupados na amostra;
- em alguns municípios, as áreas de ponderação foram definidas considerando suas divisões administrativas, sempre respeitando o critério de tamanho mínimo; alguns municípios tiveram apenas 2 áreas definidas: uma considerando todos os setores do distrito-sede e outra considerando todos os setores dos demais distritos; em outros municípios, cujos distritos possuem tamanho que feriam o critério de tamanho mínimo, também foram definidas duas áreas: uma constituída por todos os seus setores urbanos e outra por todos os seus setores rurais, mesmo que isso significasse setores não contíguos;

- para um conjunto de municípios grandes em termos de população, foi feita uma consulta aos órgãos de planejamento municipal para que as áreas de ponderação fossem definidas em conjunto. Nesses municípios, também foram considerados os critérios de tamanho mínimo e de contiguidade do conjunto de setores para a definição das áreas de ponderação; e
- os municípios que não se enquadraram nas 4 situações acima tiveram suas áreas de ponderação definidas automaticamente, usando uma metodologia de agregação de setores, implementada por meio de um sistema computacional especialmente desenvolvido, que faz uso de informações georreferenciadas; essa metodologia considera os critérios de tamanho mínimo, vizinhança entre os setores e a homogeneidade dos setores em relação a um conjunto de características conhecidas para o universo no nível dos setores. Entre as 15 variáveis utilizadas constava, por exemplo: rendimento médio dos responsáveis pelos domicílios no setor, número médio de pessoas por domicílio particular permanente, proporção de domicílios particulares permanentes ligados à rede geral de água, média de anos de estudo dos responsáveis por domicílios. Para uma descrição detalhada do procedimento, ver Silva, Matzenbacher e Cortez (2002).

A divulgação dos resultados da amostra nos diversos formatos, publicações de tabelas, CD-ROM, microdados, dados agregados em nível de área de ponderação e em nível de município, contém em sua documentação as informações para a compreensão de como resultaram as áreas de ponderação consideradas durante o processo de expansão da amostra. Essa documentação inclui uma relação dos 484 municípios que tiveram mais de uma área de ponderação com informações sobre o número de suas áreas. Os demais 5 023 municípios tiveram apenas uma área de ponderação. Além disso, inclui um arquivo com informações básicas sobre cada uma das 9.336 áreas de ponderação, a saber: código da área de ponderação; tipo da área; número de setores; número de domicílios particulares ocupados na amostra; número de pessoas no universo; e uma descrição da composição geográfica da área de ponderação.

Outro arquivo que faz parte da documentação relacionada com a definição das áreas de ponderação é o que indica o código da área de ponderação (13 posições) para cada setor censitário da base geográfica do Censo 2000, identificado também pelo seu código (15 posições).

A tabela 12.59 ao final deste item apresenta o número de municípios total e por tipo de área que contém, por Unidade da Federação.

12.5.3.3 Estimação de totais, médias e razões

As estimações de totais para domínios de interesse, como por exemplo, as células de uma tabela, são feitas, utilizando-se, para cada unidade (pessoa, família ou domicílio), o peso correspondente, que foi determinado para cada domicílio da amostra. Esse mesmo peso foi atribuído a cada pessoa moradora e a cada família do domicílio. Assim, para estimar o total de uma característica utiliza-se o estimador $\hat{Y}$ definido por:

$$\hat{Y} = \sum_{i=1}^{n} p_i \, y_i$$

onde:

p_i é o peso associado à i-ésima unidade da amostra no domínio em questão;

y_i é o valor de y associado à i-ésima unidade da amostra no domínio;

n é o número de unidades na amostra do domínio em questão.

Dessa forma, é possível calcular estimativas para quaisquer variáveis investigadas no censo, independente de serem de pessoas, famílias ou domicílios.

Os pesos calculados com a metodologia adotada não são necessariamente inteiros e não devem ser substituídos por pesos inteiros para não provocar a quebra na consistência das restrições efetivamente utilizadas no ajuste no modelo. O uso de pesos fracionários preserva o método de expansão da amostra, produz resultados mais precisos do ponto de vista estatístico. Assim, para o cálculo das estimativas das tabelas de divulgação do censo foi utilizado o peso fracionário com 8 casas decimais, sendo, então, arredondadas as estimativas resultantes.

Para obter consistência com as tabelas de divulgação, é necessário que as estimativas sejam calculadas em cada célula básica da tabela e as linhas e colunas de totais e subtotais sejam obtidas por soma das estimativas básicas correspondentes, após terem sido arredondadas. Uma conseqüência desse procedimento é que os totais de uma mesma característica podem diferir ligeiramente de uma tabela para outra, em função do arredondamento das parcelas em cada tabela.

12.5.3.4 Estimação de erros amostrais ou avaliação da orecisão das estimativas

As conclusões de uma pesquisa por amostra devem ser apoiadas nas estimativas produzidas. Essas, por sua vez, embutem um erro amostral que deve situar-se dentro de um nível de confiança fixado. Assim, a avaliação dos erros amostrais é um ponto fundamental, pois dele decorre o grau de confiança nas conclusões analíticas obtidas. Para cada estimativa derivada da pesquisa, é possível obter uma medida de precisão que auxilia na análise e interpretação dos dados resultantes da pesquisa.

Os erros amostrais podem ser avaliados através das estimativas dos coeficientes de variação ou dos erros padrão calculadas a partir das estimativas das variâncias.

Embora seja possível estimar os erros amostrais de acordo com a metodologia usada na obtenção dos pesos, o método direto é bastante complexo (SÄRNDAL; SWENSSON; WRETMAN, c1992). Sugere-se, então, um método simples e rápido para obtenção de uma aproximação do erro padrão da estimativa, que pode ser usado para a construção de intervalos com níveis de confiança fixados. Como a amostra usada no Censo Demográfico 2000 é bastante grande e os domicílios se distribuem de forma aleatória dentro de cada setor censitário, pode-se aproximar o cálculo do erro padrão, segundo Cochran (1977), supondo que o esquema de seleção da amostra foi de amostragem aleatória simples sem reposição. Dessa maneira, um estimador do erro padrão de um estimador de total de uma característica y, representado por $\hat{Y}$, é dado por:

$$ep(\hat{Y}) = \sqrt{\frac{(1-f)}{f} N s^2(y)}$$

onde:

$ep(\hat{Y})$ é o erro padrão do estimador de total, $\hat{Y}$, para o domínio em questão;

f é a fração efetiva de amostragem observada no domínio em questão;

N é o total de unidades da população no domínio em questão;

$s^2(y) = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$ é a variância amostral para o domínio em questão;

$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$ é a média amostral no domínio em questão;

y_i é o valor da característica y na i-ésima unidade da amostra no domínio;

n é o total de unidades da amostra no domínio em questão.

A divulgação dos resultados da amostra nos diversos formatos, publicações de tabelas, CD-ROM, microdados, dados agregados em nível de área de ponderação e em nível de município, contém em sua documentação um arquivo ou tabela com as frações amostrais, em porcentagem, efetivamente observadas para os domínios Brasil, Grandes Regiões, Unidades da Federação, Mesorregiões, Microrregiões e Municípios. A Tabela 12.57, a seguir, mostra o número de municípios com cada uma das duas frações efetiva aplicadas, por tamanho do município, medido em população recenseada no Censo 2000.

Tabela 12.58 - Número de municípios por fração amostral efetiva de domicílios, segundo a classe de tamanho populacional

Classes de população em 2000	Total	Fração amostral efetiva de domicílios	
		até 15%	+ de 15%
Total	5 507	2 020	3 487
Até 15 000 habitantes	3 540	114	3 426
Mais de 15 000 habitantes	1 967	1 906	61

Fonte: IBGE, Censo Demográfico 2000.

Como a maior parte das estimativas derivadas das informações coletadas na amostra do Censo Demográfico 2000 é proveniente de variáveis categóricas, para as quais y_i assume somente os valores 0 (se a unidade não pertence à categoria em questão), ou 1 (se a unidade pertence à categoria em questão), a expressão do estimador $ep(\hat{Y})$ reduz-se a:

$$ep(\hat{Y}) = \sqrt{\frac{(1-f)\hat{Y}(N-\hat{Y})}{Nf-1}}$$

Na Tabela 12.58, são apresentados valores de erros padrão calculados para alguns valores de estimativas de características de pessoas e domicílios para o Brasil.

O erro padrão é utilizado para construir intervalos de confiança que conterão o valor do total populacional[7], y_i com uma certa probabilidade decorrente do nível de confiança desejado na tomada de decisão, ou seja,

$$P[\hat{Y} - z_{\alpha/2}ep(\hat{Y}) < Y < \hat{Y} + z_{\alpha/2}ep(\hat{Y})] = 1-\alpha$$

onde:

α é o nível de significância e (1 - a) é o nível de confiança;

$z_{\alpha/2}$ é a abscissa da distribuição Normal padrão com área $\alpha/2$ à sua direita.

Assim, para um nível de confiança de 95%, tem-se $z_{\alpha/2} = 1{,}96$ e o intervalo de confiança é dado por:

$$[\hat{Y} - 1{,}96ep(\hat{Y});\ \hat{Y} + 1{,}96ep(\hat{Y})]$$

Pela Tabela 12.58, caso haja interesse em estimar um total de uma característica relativa às pessoas e essa estimativa para Brasil seja da ordem de 10 000 000, vê-se que seu erro padrão seria da ordem de 8 445. Portanto, de acordo com as fórmulas anteriores, um intervalo de 95% de confiança para o total da característica de interesse será dado por [9 983 448; 10 016 552]. Em termos percentuais, pode-se dizer que a estimativa da característica desejada é 10 000 000, com uma margem de erro relativo de 0,17% para cima ou para baixo.

Na prática, um intervalo de confiança de 95%, por exemplo, indica que, em cada 100 amostras selecionadas com o mesmo desenho, 95 produzirão estimativas cujo intervalo de confiança conterá o valor verdadeiro da população e em apenas 5 amostras este valor estará fora do intervalo de confiança.

Naturalmente, quanto maior o nível de confiança, maior será a amplitude do intervalo de confiança. A decisão sobre o nível de confiança decorre do grau de precisão que o usuário necessita em seu trabalho analítico.

A divulgação dos resultados da amostra nos diversos formatos, publicações de tabelas, CD-ROM, microdados, dados agregados em nível de área de ponderação e em nível de município, contém em sua documentação tabelas equivalentes à Tabela 12.58, para outros níveis geográficos, a saber, as 5 Grandes Regiões e as 27 Unidades da Federação:

[7] O valor da população é, de um modo geral, desconhecido, exceto para as características investigadas censitariamente.

Tabela 12.59 - Brasil - Erro padrão aproximado para alguns tamanhos de estimativas para características de pessoas e domicílios

Características de pessoas		Características de domicílios	
Tamanho da estimativa	Erro padrão aproximado	Tamanho da estimativa	Erro padrão aproximado
100	28	100	28
500	62	500	62
1 000	87	1 000	87
2 000	123	2 000	123
5 000	195	5 000	195
10 000	275	10 000	275
20 000	389	20 000	389
50 000	615	50 000	615
100 000	870	100 000	870
150 000	1 066	150 000	1 064
200 000	1 230	200 000	1 228
500 000	1 944	250 000	1 373
1 000 000	2 745	500 000	1 936
2 000 000	3 870	1 000 000	2 722
3 000 000	4 726	2 000 000	3 807
4 000 000	5 440	3 000 000	4 608
5 000 000	6 064	4 000 000	5 258
6 000 000	6 623	5 000 000	5 808
7 000 000	7 132	6 000 000	6 283
8 000 000	7 601	7 000 000	6 700
9 000 000	8 037	8 000 000	7 069
10 000 000	8 445	9 000 000	7 397
15 000 000	10 180	10 000 000	7 690
20 000 000	11 563	15 000 000	8 730
30 000 000	13 681	20 000 000	9 217
40 000 000	15 222	25 000 000	9 240
50 000 000	16 350	30 000 000	8 802
100 000 000	17 650	35 000 000	7 826
120 000 000	16 331	40 000 000	6 057
130 000 000	15 196	45 507 516	0
140 000 000	13 645		
150 000 000	11 513		
160 000 000	8 365		
169 799 170	0		

Fonte: IBGE, Censo Demográfico 2000.

Tabela 12.60 - Número de municípios total e por tipo de área que contém, por Unidade da Federação

Unidade da Federação	Quesito							
	Número de municípios existentes	Município	usuário	Distrito	Distrito-sede + Ag. Distritos[1]	Urbana + rural	Subdistrito + Ag. Subdistritos	Municípios feitos no skater
Total	**5507**	**5023**	**69**	**9**	**79**	**128**	**1**	**199**
Rondônia	52	46	1	1	0	4	0	0
Acre	22	19	0	0	0	2	0	1
Amazonas	62	57	1	0	0	4	0	0
Roraima	15	14	1	0	0	0	0	0
Pará	143	126	3	0	3	10	0	1
Amapá	16	15	1	0	0	0	0	0
Tocantins	139	137	1	0	0	0	0	1
Maranhão	217	201	3	1	0	12	0	0
Piauí	221	217	1	0	0	2	0	1
Ceará	184	147	3	0	25	8	0	1
Rio Grande do Norte	166	159	2	0	1	3	0	1
Paraíba	223	216	2	0	0	2	0	3
Pernambuco	185	155	5	0	6	13	0	6
Alagoas	101	94	1	0	0	6	0	0
Sergipe	75	69	1	0	0	4	0	1
Bahia	415	375	5	1	8	21	0	5
Minas Gerais	853	806	1	2	9	6	0	29
Espírito Santo	77	61	3	0	7	4	0	2
Rio de Janeiro	91	54	8	0	11	2	1	16
São Paulo	645	553	10	2	3	7	0	70
Paraná	399	370	3	0	2	6	0	18
Santa Catarina	293	275	3	2	1	4	0	8
Rio Grande do Sul	467	432	3	0	3	5	0	24
Mato Grande do Sul	77	72	1	0	0	2	0	2
Mato Grosso	126	122	2	0	0	1	0	1
Goiás	242	231	3	0	0	0	0	8
Distrito Federal	1	0	1	0	0	0	0	0

Fonte: IBGE, Censo Demográfico 2000.

(1) De fato, são 78 municípios nessa situação e um município (Queimados no Rio de Janeiro) que teve suas áreas definidas ou como subdistritos inteiros ou como agregados de subdistritos. Esse município foi classificado nessa categoria por ter sido o único caso nesse formato.

12.6 Tabulação dos dados

Os procedimentos de tabulação dos dados constituem a última etapa do processo de apuração das informações. No Censo Demográfico 2000, a tabulação teve início logo após o encerramento da coleta, com os Resultados Preliminares, continuou com a Sinopse Preliminar, alcançando, mais tarde, os resultados referentes ao Conjunto Universo e ao Questionário da Amostra.

Com exceção das tabulações dos Resultados Preliminares e da Sinopse Preliminar, que tiveram como base os dados do SIGC, ainda não submetidos ao processo de crítica, as demais tiveram seu início imediatamente após a realização desse trabalho de depuração de erros.

Independentemente do tipo de publicação, o processo de tabulação tem como referência uma proposta de plano tabular, discutida no âmbito do Comitê do Censo Demográfico 2000 e com os principais usuários externos das informações censitárias.

O início efetivo do processo de tabulação aconteceu com a confecção das molduras das tabelas, que corresponde à elaboração dos textos referentes aos títulos e rodapés, bem como os que constituem as indicações das colunas de dados e coluna indicadora de cada tabela. Esse trabalho foi desenvolvido pelo Centro de Documentação e Disseminação de Informações – CDDI, obedecendo-se às *Normas de apresentação tabular* (1993), para, mais tarde, receber a aprovação da Coordenação de Planejamento e Organização do Censo Demográfico 2000 – CPO.

O passo seguinte foi a elaboração da seleção de variáveis, ocasião em que técnicos da Diretoria de Pesquisas - DPE organizaram, para cada uma das tabelas do plano, as formas de obtenção dos seus dados, identificando o manejo dos códigos ou dos valores das variáveis coletadas nos questionários ou criadas durante o processo de crítica das informações.

O trabalho continuou com a entrega das molduras e respectivas seleções de variáveis a dois grupos de técnicos, um da Diretoria de Informática - DI e outro da DPE, para a obtenção dos dados. Para esse trabalho de dupla programação, a DI utilizou o Sistema Pegasus e a DPE o Sistema SAS.

Sistema Pegasus de tabulação

O Pegasus, sistema de tabulação utilizado no Censo 2000, foi desenvolvido em Visual Basic- VB-v6, FrontPage e Opus, sendo ativado via Internet Explorer 4.0. Assim, os diálogos foram implementados em VB-v6, as páginas *Web* desenvolvidas no FrontPage, as rotinas de tabulação em Opus e a visualização das tabelas em Excel 7.0. O sistema é composto por:

- banco de dados em um servidor RISC, que armazena a definição dos planos de tabulação, os dicionários e também as matrizes com os valores tabulados, que são visualizados, via browser, de qualquer microcomputador ligado à rede do IBGE com a configuração necessária para acesso ao Pegasus – Windows 98, Internet Explorer 4.0, MS Office 97; e

- rotinas com os diálogos, residentes em uma máquina Windows NT com um servidor *Web*, e que são instaladas na máquina do operador automaticamente pelo Internet Explorer, quando conectado ao sistema pela primeira vez;

O funcionamento do Pegasus se inicia quando a página *Web* de apresentação é chamada por um operador. Nela existem os links para funções de auxílio, para os manuais e para o sistema

O processo de tabulação dos microdados é bastante rápido, feito por um programa OPUS, residente em um servidor Unix que, para cada plano tabular, faz a leitura dos dados uma única vez.

Sistema de conferência

Um sistema de conferência, operando *on-line*, baseado em macros do Excel, implementava a comparação das tabelas geradas pelos Sistemas Pegasus e SAS. Nessa oportunidade, todos os elementos das tabelas eram cotejados, sendo qualquer diferença, entre seus textos ou valores, apontada pela rotina ao técnico operador.

Elegia-se para teste uma Unidade da Federação, para a qual todas as tabelas do plano passavam pelo processo de conferência. Somente após sanados as divergências, era gerado o plano tabular para todas as outras unidades e demais níveis de divulgação.

Deve-se ressaltar a importância que o sistema de conferência implementado representou para a velocidade e segurança de todo o processo de tabulação dos dados do censo, sendo que a rotina foi aperfeiçoada ao longo dos trabalhos para as diversas divulgações realizadas, tendo o início desse processo acontecido com a Sinopse Preliminar.

Referências

ALBIERI, S. *Apresentação da precisão de estimativas nas tabelas de pesquisas por amostragem do IBGE*. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1999a.

_____. *A ausência de resposta em pesquisas:* uma aplicação de métodos de imputação. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 1992. 138 p. (Informes de matemática. Série D-048/92). Dissertação de mestrado apresentada em 1989.

_____. *Considerações preliminares para o planejamento da amostra para a tabulação avançada do censo demográfico 2000*. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 1999b.

_____. *Nota técnica sobre a definição do tamanho das áreas de ponderação do censo demográfico 2000*. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2003. 7 p.

_____; DIAS, A. J. R.; MENEZES, A. C. F.; GREEN, A. P. L. *Controle de qualidade da captura de dados do censo 2000:* especificações para o planejamento das amostras: (2ª versão). Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2001.

_____; MARTELOTTE, M. C.; DUARTE, R. P. N. *Estudos para subsidiar o planejamento da amostra para a tabulação avançada do censo demográfico 2000*. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2000.

BANKIER, M. D. *Two step generalized least squares estimation*. Ottawa: Statistics Canada, 1990. 66 p.

BANKIER, M. D. et al. Imputing numeric and qualitative variables simultaneously. *Proceedings of the Survey Research Methods Section*, Baltimore, 1996, p.90-99, [1996?].

BARBOSA, D. M. R.; SILVA, A. do N. *NIM - new imputation methodology.* Rio de Janeiro: IBGE, Diretoria de Informática, 2002. (Nota Técnica, 02/02).

BRASS, W. *Seminário sobre métodos para medir variables demográficas:* fecundidad y mortalidad, 1971, São José. San José: CELADE, 1973. 146 p. (Serie DS CELADE, n. 9).

_____ et al. *The demography of tropical Africa.* Princeton: Princeton University Press, 1968.

_____. Estimating mortality from deficient registration data. In: _____. *Methods for estimating fertility and mortality from limited and defective data*. Chapel Hill: University of North Carolina, 1975. p.117-123. (Laboratories for Population Statistics. An occasional publication).

BRAVO, P. C. Elementos de controle estatístico de qualidade. In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 6., 1984, Rio de Janeiro. *Atas...* Rio de Janeiro: UFRJ, 1984. 585 p.

BREIMAN, L. et al. *Classification and regression trees*. Belmont, Calif.: Wadsworth International Group, c1984. (The Wadsworth statistics/probability series).

CAMISA, Z. Fecundidad y nupcialidad. In: ENCUESTA demográfica nacional de Honduras. [Tegucigalpa]: Dirección General de Estadística y Censos de Honduras; Santiago de Chile: CELADE, 1975. v. 3. (Série A. CELADE, n. 129).

CENSO DEMOGRÁFICO 2000: características da população e dos domicílios: resultados do universo. Rio de Janeiro: IBGE, 2001.

_____: fecundidade e mortalidade infantil: resultados preliminares da amostra. Rio de Janeiro: IBGE, 2002. 21 p.

_____: resultados preliminares. Rio de Janeiro: IBGE, 2000. 172 p.

CENSO demográfico 2000: manual do recenseador CD 1.09. Rio de Janeiro: IBGE, 2000. 151 p.

CENTRY user's guide: IMPS version 3.1. Washington, D.C.: Bureau of the Census International Systems Team, 1995.

COALE, A. J.; DEMENY, P. *Regional model life tables and stable populations*. Princeton, N.J.: Princeton University Press, 1966. 871 p.

_____; TRUSSELL, J. Estimating the time to which Brass estimates Apply, annex 1 to Samuel H. Preston and Alberto Palloni, fine-tuning Brass-type mortality estimates with data on ages of surviving children. *Population Bulletin of the United Nations*, New York, n. 10, p. 87-89, 1977.

_____. Model fertility schedules: variations in the age structure of childbearing in human populations. *Population Index*, Princeton, v. 40, n. 2, p. 185-257, 1974.

COCHRAN, W. G. *Sampling techniques*. 3rd ed. New York: Wiley, c1977. 428 p.

CONCOR user's guide: IMPS version 3.1. Washington, D.C.: Bureau of the Census International Systems Team, 1995.

CORREÇÃO de declaração de dados de fecundidade e mortalidade. [Rio de Janeiro: IBGE, 2002]. 4 p.

DEFINIÇÕES necessárias à implementação da apuração centralizada dos dados referentes ao questionário da amostra (CD 1.02) no censo demográfico de 2000. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2003. 78 p.

DIAS, A. J. R. *Aspectos de amostragem do censo demográfico de 2000*. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2002. 17 p.

ESQUEMA de ordenação lógica das pessoas no questionário básico - censo 2000. Rio de Janeiro: IBGE, 2001.

FEENEY, G. Estimating infant mortality rates from child survivorship data by age of mother. *Asian and Pacific Census Newsletter*, v. 3, n. 2, p. 12-16, Nov. 1976.

_____. Estimating infant mortality trends from child survivorship data. *Population Studies*, London, v. 34, n. 1, p. 109-128, Mar. 1980.

FRIAS, L. A. de M.; RODRIGUES, P. *Brasil*: tábuas-modelo de mortalidade e populações estáveis. Rio de Janeiro: IBGE, 1981. 149 p. (Estudos e pesquisas, n. 10).

GENERALIZED *estimation system*: version 4.0: help guide. Ottawa: Statistics Canada, 1998.

GUIMARÃES, N. R. *Controle de qualidade do censo 2000:* análise final dos processos de reconhecimento e verificação. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2001. 54 p.

_____; DIAS, A.; ALBIERI, S. *Censo demográfico 2000:* controle de qualidade do processo de captura dos dados. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2001.

HRADESKY, J. L. *Productivity and quality improvement*: a practical guide to implementing statistical process control. New York: McGraw-Hill, 1988. 243 p.

LEHMANN, E. L. *Nonparametrics:* statistical methods based on ranks. San Francisco: Holden-Day, c1975. 457 p.

LITTLE, R. J. A.; RUBIN, D. B. *Statistical analysis with missing data*. New York: Wiley, 1987.

MANSOLDO, H. M.; SILVA, A. C. C. M. *Codificação automática/assistida*: uma proposta para o censo demográfico 2000. Rio de Janeiro: IBGE, Diretoria de Informática, 1997.

MANUAL X: indirect techniques for demographic estimation. New York: United Nations, Department of International Economic and Social Affairs, 1983. 304 p. (Population studies, n. 81).

MANUAL de crítica: sistema IMPS. Rio de Janeiro: IBGE, 2002.

MONTGOMERY, D. C. *Introduction to statistical quality control*. 3rd ed. New York: Wiley, c1996. Várias paginações.

NORMAS de apresentação tabular. 3. ed. Rio de Janeiro: IBGE, 1993. 62 p.

OLIVEIRA, J. de C. *Fecundidade e nupcialidade no Brasil e nos estados de São Paulo e Rio Grande do Norte:* tendências passadas e perspectivas. Rio de Janeiro: IBGE, 1991. 133 p.

PESSOA, D. G. C.; MOREIRA, G. G.; SANTOS, A. R. *Imputação de rendimentos no questionário da amostra do censo demográfico 2000*. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2003. 17 p.

_____; SANTOS, A. R. *Imputação de rendimento dos responsáveis por domicílios – conjunto universo do censo demográfico 2000*. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2003. 15 p.

_____; SILVA, P. L. N. *Análise de dados amostrais complexos*. São Paulo: Associação Brasileira de Estatística, 1998.

PLANO de análise da correção automática e elementos de apoio para a análise da formação do lote, sistema DIA, CD 1.02 - questionário da amostra. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2002. 37 p.

PROCEDIMENTOS a serem implementados para a execução da crítica (NIM e IMPS), questionário da amostra/censo 2000. Rio de Janeiro: IBGE, 2001.

RELATÓRIO sobre a experiência de tratamento automático de críticas entre registros, com vistas ao censo demográfico do ano 2000. [Rio de Janeiro]: IBGE, Diretoria de Pesquisas, 1999. 4 p.

RUBIO, E.; CRIADO, I. V. *Sistema DIA:* sistema de detección e imputación automática de errores para datos cualitativos. Madrid: Instituto Nacional de Estadística, 1988. v. 1: DIA : descripción del sistema.

SÄRNDAL, C. E.; SWENSSON, B.; WRETMAN, J. *Model assisted survey sampling.* New York: Springer-Verlag, c1992. 694 p.

SAS procedures guide: version 6. 3rd ed. Cary, NC: SAS Institute, c1990. 705 p.

SHRYOCK, H. S. et al. *The methods and materials of demography*. [Washington, D.C.]: Bureau of the Census, 1971. v.1.

SILVA, A. N. *Algumas considerações sobre o uso do NIM no censo demográfico 2000*. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2003. 18 p.

_____; CORTEZ, B. F. *Censo demográfico 2000:* formação de lotes para a crítica de estrutura dos questionários da amostra – CD 1.02. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2000. 11 p.

_____; MATZENBACHER, L. A.; CORTEZ, B. F. *Processamento das áreas de expansão e disseminação da amostra no censo demográfico 2000*. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2002. 28 p.

SILVA, L. F.; BIANCHINI, Z. M. *A redução da amostra e a utilização de duas frações amostrais no censo demográfico de 1990*. Rio de Janeiro: IBGE, 1990. 49 p. (Textos para discussão, n. 33).

SILVA, P. L. do N.; BIANCHINI, Z. M.; ALBIERI, S. *Uma proposta de metodologia para a expansão da amostra do censo demográfico de 1991*. Rio de Janeiro: IBGE, 1993. 106 p. (Textos para discussão, n. 62).

_____; PESSOA, D. G. C. *Estimando a precisão das estimativas das taxas de mortalidade obtidas a partir da PNAD*. Rio de Janeiro: IBGE, Diretoria de Pesquisas, 2002. 24 p.

SINOPSE preliminar do censo demográfico 2000. Rio de Janeiro: IBGE, v. 7, 2000. Acompanha 1 CD-ROM.

TABULAÇÃO avançada do censo demográfico 2000: resultados preliminares da amostra. Rio de Janeiro: IBGE, 2002.

TRUSSELL, T. J. A re-estimation of the multiplying factors of the Brass technique for determining childhood survi-vors-hip rates. *Population Studies*, London, v. 19, n. 3, p. 97-107, 1975.

VENABLES, W. N.; RIPLEY, B. D. *Modern applied statistics with S-Plus.* New York: Springer, 1994.

WERKEMA, M. C. C. *Avaliação da qualidade de medidas*. Belo Horizonte: UFMG, Escola de Engenharia: Fundação Christiano Ottoni, 1996. 101 p. (Ferramentas da qualidade, 13).

_____. *Ferramentas estatísticas básicas para o gerenciamento de processos*. Belo Horizonte: UFMG, Escola de Engenharia: Fundação Christiano Ottoni, 1995. 384 p.

13 Disseminação

Concluídas as fases de planejamento, coleta e apuração, a etapa seguinte, de disseminação, colocou à disposição da sociedade os resultados do Censo Demográfico 2000. Nesta etapa, as informações resultantes da última operação censitária do século foram levadas ao conhecimento da população, dos órgãos governamentais, das empresas privadas e de todos os outros segmentos e classes. A cada lançamento de resultados, a imprensa era mobilizada, e os dados do Censo 2000 divulgados, de forma ampla, nos diversos meios de comunicação.

A disseminação dos resultados do Censo foi realizada por meio de publicações impressas, CD-ROMs, mapas, internet, entre outros tipos de produtos e serviços, especificados neste capítulo. Dessa forma, o IBGE reforçou o cumprimento de sua missão de retratar o Brasil com informações necessárias ao conhecimento de sua realidade e ao exercício da cidadania.

13.1 O Censo e a mídia

A realização do Censo 2000, pelas suas dimensões e pela abrangência da operação, foi um espaço privilegiado que permitiu a formulação e a consolidação de um novo modo de levar à sociedade os resultados dos trabalhos realizados pelo IBGE.

A definição das políticas de lançamento da pesquisa e da sustentação da operação censitária, destinadas a informar à população sobre a importância do Censo, envolveu a Presidência, o Centro de Documentação e Disseminação de Informações – CDDI e a Coordenação Operacional dos Censos – COC. Esse trabalho

tinha objetivo preciso: facilitar o acesso dos Recenseadores aos domicílios, com vistas à realização das entrevistas.

Através da agência de publicidade licitada para o Censo 2000[1], foi contratada uma empresa específica para assessoria de imprensa – a BR Comunicações. A contratação de uma empresa para o trabalho de relação com a mídia foi uma decisão oportuna e seus resultados podem ser medidos pelo espaço alcançado pelo Censo na mídia nacional, regional e local. A presença do Presidente da Instituição nos programas de maior audiência da televisão brasileira, a divulgação de reportagens e as entrevistas nos diversos meios de comunicação, em todo o País, foram importantes para sustentar a mídia ao longo de todo o período da coleta.

Os arquivos que documentam este trabalho mostram a sua dimensão, com destaque para a integração entre as Unidades Regionais e os meios de comunicação. Graças a isso foi possível direcionar o trabalho jornalístico de divulgação da importância da participação da população, visando facilitar o trabalho dos Recenseadores. Como exemplo, no período final da coleta, diante da dificuldade de acesso aos domicílios, as TVs divulgaram matérias, enfatizando a importância da participação da população, em atender ao Recenseador.

Dois fatos devem ser ressaltados nesse processo:

1. na medida em que eram solicitados, os responsáveis pelos meios de comunicação conscientizavam-se da importância do Censo para o País, o que, certamente, abriu as portas para a exposição que os números tiveram depois;

2. este trabalho de confiança recíproca facilitou a busca de uma nova política de comunicação para a divulgação das pesquisas do IBGE, estabelecendo as bases para uma parceria informal da Instituição com os veículos de comunicação de todo o País. Esta relação, pelas suas características, contribuiu para tornar mais visível a competência e a credibilidade do IBGE.

Concluídas as fases de lançamento e de sustentação durante o período de coleta de informações, no qual atuou a BR Comunicações, o processo de divulgação dos resultados, nas suas diferentes etapas, passou a ser de responsabilidade da Coordenação de Comunicação Social do IBGE.

A criação do sistema de embargo – que consistia em liberar para os meios de comunicação os resultados e o conteúdo técnico antecipadamente, mas com o compromisso de só serem divulgados na data estabelecida pelo IBGE – facilitou e permitiu aos veículos de comunicação maior disponibilidade aos números e melhor conhecimento deles, através do contato e da consulta aos técnicos responsáveis pelo processo de elaboração do material a ser divulgado.

O resultado desse esforço permitiu que o Censo tivesse, nos meios de comunicação, a repercussão, o espaço e o tempo que teve, em todas as diversas fases e etapas de divulgação, desde os *Resultados preliminares* até a lançamento dos volumes temáticos – *Características gerais da população, Migração e deslocamento, Educação, Trabalho e rendimento, Nupcialidade e fecundidade* e *Famílias e domicílios* – e do *Atlas do censo demográfico 2000*.

Mesmo assim, alguns problemas aconteceram. Não houve, por exemplo, melhor aproveitamento da divulgação dos resultados definitivos da amostra. A premência de tempo, determinada pelo cumprimento da meta, não permitiu

[1] Ver capítulo 4, item 4.4.

melhor divulgação do material. Da mesma forma, a divulgação das informações relativas aos municípios com mais de 100 mil habitantes, careceu de maior cobertura dos meios de comunicação regional e local.

Do ponto de vista externo, houve somente um problema. Uma revista semanal de circulação nacional contestou o Censo na sua natureza – a denúncia: a população não foi contada como devia ter sido. O IBGE respondeu como devia: refutou os argumentos da revista com ampla gama de informações e não se furtou a qualquer explicação. A denúncia da revista caiu no vazio. Esgotou-se nela mesma. Todos os demais meios de comunicação do País, não tomaram conhecimento e nem se preocuparam em consultar o IBGE – simplesmente ignoraram o assunto.

A documentação existente sobre a divulgação do Censo em todas as suas fases nos meios de comunicação é rica – um arquivo de *clipping* com mais de duas mil reportagens, artigos e editoriais da imprensa escrita, imagens, vídeos. Além do documentário *O País é este*, produzido pelo cineasta Zelito Viana, a partir dos resultados do Censo 2000.

13.2 Produtos e serviços

A divulgação de resultados do Censo Demográfico 2000 procurou levar a cada segmento de usuários as mídias mais apropriadas, fazendo uso de maneira intensiva das tecnologias digitais.

13.2.1 Publicações impressas

Tradicionalmente em Censos anteriores, a divulgação do plano tabular ocupava mais de uma centena de volumes impressos, oferecendo dificuldade de manuseio e de guarda. A solução encontrada e já testada com sucesso na Contagem da População 1996 foi a de produzir cada publicação em papel acompanhada de um CD-ROM.

O volume impresso apresenta as tabelas básicas, acrescidas de comentários, gráficos e cartogramas, a fim de tornar a publicação mais atraente para o leitor. Já o CD-ROM encartado contém a publicação impressa em pdf e as tabelas em formato de planilha eletrônica, acrescida por um conjunto adicional de tabelas e de níveis geográficos. Dessa forma, o usuário tem acesso a um produto mais completo e versátil a custos menores.

A seguir estão listadas as publicações impressas contendo resultados do Censo 2000, por ordem cronológica de lançamento.

2000

Censo demográfico 2000: resultados preliminares

Divulga os primeiros resultados do Censo Demográfico 2000, relativamente aos 5 507 municípios brasileiros instalados em 1º de agosto de 2000, data de referência da pesquisa.

Apresenta, em três tabelas, o total da população residente por sexo e situação do domicílio e a taxa média de crescimento anual, incluindo dados comparativos com os recenseamentos realizados em 1980 e 1991 e com a pesquisa Contagem da População, efetuada em 1996, segundo as grandes regiões, unidades da federação e municípios.

Esses resultados têm origem na caderneta do setor, estando sujeitos a alterações quando da apuração dos questionários preenchidos nos domicílios.

Inclui notas metodológicas sobre alguns aspectos da coleta, cartograma do Brasil e das unidades da federação, além de tabela com os municípios instalados a partir de 1º de janeiro de 2001 e os municípios de origem, com suas respectivas populações. No CD-ROM que acompanha a publicação também foram incluídas informações desagregadas para regiões metropolitanas.

2001

Sinopse preliminar do censo demográfico 2000

Apresenta os dados preliminares do Censo Demográfico 2000, em um único volume, com informações sobre a população para o total do Brasil e as regiões metropolitanas, incluindo tabelas comparativas entre os censos realizados anteriormente, bem como sobre a população residente e os domicílios, por unidade da federação e municípios.

Este livro é o único que divulga dados sobre os domicílios segundo sua espécie: particular ocupado e não-ocupado (fechado, de uso ocasional e vago) e coletivos. É também a única publicação de resultados do Censo 2000 que divulga os dados da população residente para as diversas situações urbanas e rurais existentes, como área urbanizada, área não urbanizada, área urbana isolada, aglomerados rurais de extensão urbana, povoado, núcleo e outros aglomerados.

Esses resultados têm origem na caderneta do setor censitário, sujeitos a alterações quando da apuração dos questionários preenchidos nos domicílios. Cada unidade da federação é acompanhada de dois mapas, que apresentam a distribuição da população residente urbana e rural e a taxa de crescimento da população.

A publicação apresenta uma análise sobre a dinâmica da população brasileira, avaliando o crescimento, a concentração e a razão de sexos da população, como também a condição de ocupação e média de moradores por domicílios. Inclui notas metodológicas da pesquisa, além de tabela com os municípios instalados a partir de janeiro de 2001 e os municípios de origem com suas respectivas populações. Além destes resultados, o CD-ROM que acompanha a publicação reúne informações por unidade da federação, detalhadas para os municípios e distritos.

Tendências demográficas: uma análise dos resultados da sinopse preliminar do censo demográfico 2000

Reúne um conjunto de dados e indicadores, apresentados sob a forma de tabelas, gráficos e mapas, que sintetizam os resultados da Sinopse Preliminar do Censo Demográfico 2000, acompanhados de uma análise sobre a evolução da população, grau de urbanização, taxa de crescimento, concentração populacional, razão de sexo, número de domicílios particulares existentes e média de moradores, com base nos resultados dos últimos censos, que evidenciam as variações no ritmo do crescimento demográfico ao longo do período.

Os resultados são divulgados para o total do país, grandes regiões, unidades da federação, regiões metropolitanas e municípios.

Censo demográfico 2000: características da população e dos domicílios: resultados do universo

Apresenta os dados definitivos do universo da pesquisa, que abrangem as características da população, das pessoas responsáveis pelos domicílios e dos domicílios e seus respectivos moradores, relativos ao total do país, grandes regiões, unidades da federação e municípios.

Inclui comentários sobre a dinâmica da população brasileira, analisando seu crescimento demográfico e sua distribuição, bem como dos responsáveis pelos domicílios, por sexo, idade e educação, e sobre a qualidade de vida da população através de suas condições habitacionais.

Apresenta, ainda, notas metodológicas da pesquisa, um anexo com a relação da população residente, segundo os municípios de origem em 2000 e os instalados em 2001, e um apêndice com a lista das 1 052 tabelas de resultados incluídas no CD-ROM que acompanha a publicação, abrangendo as informações do volume impresso e dados desagregados para mesorregiões, microrregiões, regiões metropolitanas, distritos, subdistritos e bairros, para cidades que têm delimitação legal de bairro.

2002

Perfil das mulheres responsáveis pelos domicílios no Brasil 2000

Reúne indicadores selecionados que retratam o perfil da mulher considerada como referência no domicílio, a partir dos resultados do Censo Demográfico 2000.

Apresenta a distribuição geográfica e urbano-rural do universo feminino responsável pelo domicílio, bem como seu nível de escolaridade e de rendimento e grupos de idade. Inclui informações sobre as mulheres que moram sozinhas, indicadores referentes às crianças na primeira infância que vivem sob a responsabilidade feminina e dados comparativos dos Censos Demográficos de 1991 e 2000. A publicação é acompanhada de CD-ROM que inclui todas essas informações e indicadores para os municípios brasileiros.

Tabulação avançada do censo demográfico 2000: resultados preliminares da amostra

Apresenta estimativas preliminares para um conjunto de variáveis investigadas no questionário da amostra, relativas aos temas: características gerais da população, educação, migração, nupcialidade, trabalho, famílias e domicílios.

Contém informações sobre a população residente por sexo, cor ou raça, religião, categorias de incapacidade ou deficiência física ou mental, nível educacional, movimentos migratórios, situação conjugal, ocupação, rendimento de trabalho, tamanho da família e condições habitacionais, entre outros aspectos, para o total do Brasil, grandes regiões e unidades da federação.

Inclui comentários sobre as características selecionadas em cada um dos temas e as diferenças regionais observadas, notas e anexos contendo a Classificação de Religiões, desenvolvida pelo IBGE e o Instituto Superior de Estudos da Religião - ISER, e a Classificação Nacional de Atividades Econômicas - Domiciliar.

Os resultados são estimativas baseadas nas informações de 108 989 domicílios e 423 049 pessoas, selecionados na amostra do Censo Demográfico 2000. A publicação é acompanhada de CD-ROM.

Censo demográfico 2000: fecundidade e mortalidade infantil: resultados preliminares da amostra

Apresenta uma análise das estimativas preliminares de variáveis investigadas no questionário da amostra, relativas aos temas fecundidade e mortalidade infantil. Os comentários destacam as alterações ocorridas na estrutura e nos níveis de fecundidade feminina, por grandes regiões e grupos de idade, com base

nos censos demográficos a partir de 1940, bem como os resultados das estimativas regionais da mortalidade infantil para a década de 1990. Inclui a conceituação das características investigadas. A publicação é acompanhada de CD-ROM.

Perfil dos idosos responsáveis pelos domicílios no brasil 2000

Reúne indicadores que retratam o perfil das pessoas de 60 anos ou mais de idade, especialmente daquelas responsáveis pelo domicílio, a partir dos resultados do universo do Censo Demográfico 2000.

Apresenta a distribuição geográfica desse grupo populacional, por grandes regiões, unidades da federação e municípios das capitais, destacando-se sua condição no domicílio, nível de escolaridade e rendimento mensal, bem como informações sobre os arranjos familiares nos domicílios sob a responsabilidade de idosos e as condições de saneamento básico desses domicílios.

Inclui comentários sobre os indicadores selecionados, com dados comparativos com o Censo Demográfico de 1991, e as diferenças regionais observadas. Esses resultados estão disponíveis no CD-ROM que acompanha a publicação, incluindo, também, informações sobre o perfil dos idosos responsáveis por domicílios nos 5 507 municípios instalados em 1º de agosto de 2000, data de referência da pesquisa.

Indicadores sociais municipais 2000: Brasil

Contém uma síntese dos indicadores da realidade socioeconômica dos municípios brasileiros, a partir das informações do universo do Censo Demográfico 2000.

Apresenta as principais características da população e dos domicílios, enfocando o perfil socieconômico de seus responsáveis, como rendimento e educação, bem como informações sobre as crianças de zero a seis anos de idade em condições habitacionais desfavoráveis ao seu desenvolvimento, para o Brasil e as unidades da federação. A análise dos resultados revela as modificações observadas nas condições de vida da população nesse período. Inclui a conceituação utilizada.

A construção desses indicadores, apresentados hierarquicamente segundo o porte populacional dos municípios em cada unidade da federação, permite a comparabilidade com os resultados obtidos no Censo Demográfico de 1991. A publicação é acompanhada de CD-ROM que inclui todas essas informações e os indicadores para os 5 561 municípios brasileiros existentes em 2001, além de uma análise dos resultados por unidade da federação.

Tendências demográficas: uma análise dos resultados do universo do censo demográfico 2000

Reúne um conjunto de dados e indicadores obtidos a partir dos resultados do universo do Censo Demográfico 2000, abrangendo informações sobre a distribuição territorial da população brasileira, estrutura etária, crescimento do contingente de idosos, evolução da alfabetização e perfil socioeconômico dos responsáveis pelos domicílios particulares permanentes, bem como as condições de moradia e acesso aos serviços de saneamento básico desses domicílios.

O volume relativo ao Brasil apresenta os resultados em tabelas e gráficos por grandes regiões, unidades da federação e regiões metropolitanas, incluindo análises sobre as variações e tendências observadas nesses indicadores com

base nos últimos censos. O volume relativo a cada unidade da federação contém a análise dos resultados obtidos para essa área geográfica, ilustrados através de gráficos. A publicação é acompanhada de CD-ROM.

Censo demográfico 2000: primeiros resultados da amostra

Divulga os resultados definitivos da amostra do Censo Demográfico 2000, relativos aos temas: características gerais da população, educação, migração, deslocamento, nupcialidade, fecundidade, trabalho e rendimento, famílias e domicílios.

Apresenta dados sobre a população residente por sexo, cor ou raça, religião, categorias de incapacidade ou deficiência física ou mental, anos de estudo, lugar de nascimento, deslocamentos realizados para o trabalho ou estudo, estado civil, filhos nascidos vivos, óbitos fetais e filhos sobreviventes das mulheres de 10 anos ou mais de idade, grupos ocupacionais, condição de ocupação, atividade econômica e rendimento de trabalho, bem como informações sobre o número de componentes das famílias residentes em domicílios particulares permanentes, existência de bens duráveis e infra-estrutura de saneamento básico nesses domicílios, entre outros aspectos.

A publicação está dividida em duas partes e apresenta os resultados nacionais por grandes regiões e unidades da federação. O CD-ROM que a acompanha contém informações também para as regiões metropolitanas e municípios, bem como anexos com as classificações utilizadas. Ambos incluem notas metodológicas e os conceitos e definições da pesquisa. Os resultados constituem estimativas obtidas pela expansão da amostra de 5 304 711 domicílios e 20 274 412 pessoas.

2003

Censo demográfico 2000: características gerais da população: resultados da amostra

Apresenta informações selecionadas sobre a estrutura por idade, sexo, situação do domicílio, cor ou raça, religião e níveis de incapacidade ou deficiência física ou mental da população residente no Brasil.

A publicação inclui comentários sobre a evolução das características socioeconômicas selecionadas comparativamente aos resultados obtidos no Censo Demográfico 1991, as diferenças observadas nos diferentes estratos geográficos e populacionais do país, bem como notas metodológicas e os conceitos e definições utilizados. O CD-ROM que a acompanha contém, além das informações do volume impresso, dados para grandes regiões e unidades da federação, a íntegra do plano tabular objeto de divulgação do volume *Censo Demográfico 2000: primeiros resultados da amostra*, e anexo contendo a Classificação de Religiões, desenvolvida pelo IBGE e o Instituto Superior de Estudos da Religião – ISER.

Os resultados constituem estimativas obtidas pela expansão da amostra de 5 304 711 domicílios e 20 274 412 pessoas, selecionados em todo o território nacional.

Censo demográfico 2000: migração e deslocamento: resultados da amostra

Apresenta informações selecionadas sobre os movimentos migratórios da população e seus deslocamentos, referentes ao total do país e grandes regiões.

A publicação inclui comentários sobre entradas, saídas e saldos migratórios comparativamente aos resultados obtidos no Censo Demográfico 1991 e sobre as pessoas que efetuaram deslocamentos para fora do município de resi-

dência com a finalidade de estudo ou trabalho. São apresentadas, também, as diferenças observadas nos diferentes estratos geográficos e populacionais do país, bem como notas metodológicas e os conceitos e definições utilizados.

O CD-ROM que a acompanha contém, além das informações do volume impresso, dados para unidades da federação, mesorregiões, microrregiões, e municípios, e a íntegra do plano tabular objeto de divulgação do volume *Censo Demográfico 2000: primeiros resultados da amostra*. Os resultados constituem estimativas obtidas pela expansão da amostra de 5 304 711 domicílios e 20 274 412 pessoas, selecionados em todo o território nacional.

Censo demográfico 2000: educação: resultados da amostra

Apresenta indicadores selecionados que permitem caracterizar o perfil educacional da população e os principais aspectos da democratização da educação no Brasil, para o total do país e grandes regiões.

A publicação contempla informações sobre a condição de alfabetização, freqüência à creche e à escola, anos de estudo e grau de escolarização concluído, segundo a cor ou raça, grupos de idade, situação do domicílio e níveis de ensino, com enfoque especial sobre o contingente populacional com educação superior completa, e inclui comentários analíticos de especialistas do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP. São apresentadas, também, notas metodológicas e os conceitos e definições utilizados.

O CD-ROM que a acompanha contém, além das informações do volume impresso, dados para unidades da federação, mesorregiões, microrregiões, e municípios, e a íntegra do plano tabular objeto de divulgação do volume *Censo Demográfico 2000: primeiros resultados da amostra.* Os resultados constituem estimativas obtidas pela expansão da amostra de 5 304 711 domicílios e 20 274 412 pessoas, selecionados em todo o território nacional.

Censo demográfico 2000: trabalho e rendimento: resultados da amostra

Apresenta as principais características das pessoas de 10 anos ou mais de idade, no que se refere à sua condição de atividade, à sua forma de inserção no mercado e os rendimentos auferidos do trabalho como de outras fontes.

As informações divulgadas referem-se à escolaridade, cobertura previdenciária, posição na ocupação, grupos de ocupação e de atividade econômica no trabalho principal, jornada de trabalho, nível e concentração dos rendimentos do contingente populacional em idade ativa, tanto oriundos de trabalho como de todas as fontes, e das famílias, entre outros indicadores. Os comentários enfocam as distinções entre determinados contingentes populacionais e evidenciam as diferenças entre as estruturas produtivas e nos níveis e concentrações dos rendimentos das grandes regiões do País. São apresentadas, também, notas metodológicas e os conceitos e definições utilizados.

O CD-ROM que acompanha a publicação contém, além das informações do volume impresso, dados para unidades da federação, mesorregiões, microrregiões e municípios, e a íntegra do plano tabular objeto de divulgação do volume *Censo Demográfico 2000: primeiros resultados da amostra.* Os resultados constituem estimativas obtidas pela expansão da amostra de 5 304 711 domicílios e 20 274 412 pessoas, selecionados em todo o território nacional.

Censo demográfico 2000: nupcialidade e fecundidade: resultados da amostra

Apresenta as principais características demográficas e socioeconômicas das pessoas de 10 anos ou mais de idade por estado conjugal e os fatores condicionantes da fecundidade das mulheres desse mesmo grupo etário.

O conjunto de resultados acerca da nupcialidade abrange informações sobre a condição de convivência, estado civil e natureza da união das pessoas de 10 anos ou mais de idade. No que concerne à fecundidade, são investigados o número de mulheres que tiveram filhos, filhos tidos – nascidos vivos e nascidos mortos – e filhos sobreviventes. A análise dessas informações enfoca os principais determinantes de ordem social, econômica e cultural observados em cada um dos temas, destacando-se cor ou raça, religião, nível educacional, rendimento mensal familiar, entre outros aspectos, bem como os contrastes observados nos diferentes estratos geográficos e populacionais do país.

Complementando esse conjunto de informações, é apresentado um estudo sobre a mortalidade infantil, com enfoque especial sobre as desigualdades sociais e regionais identificadas na análise desse indicador, tendo em vista sua importância para a avaliação das condições de saúde e de vida da população do país. São apresentadas, também, notas metodológicas e os conceitos e definições utilizados.

O CD-ROM que acompanha a publicação contém, além das informações do volume impresso, dados para as unidades da federação, mesorregiões, microrregiões e municípios, e a íntegra do plano tabular objeto de divulgação do volume *Censo Demográfico 2000: primeiros resultados da amostra.*

Os resultados constituem estimativas obtidas pela expansão da amostra de 5 304 711 domicílios e 20 274 412 pessoas, selecionados em todo o território nacional.

Censo demográfico 2000: famílias e domicílios: resultados da amostra

Apresenta os principais aspectos da organização das famílias bem como informações sobre as condições habitacionais da população e as características gerais do entorno dos domicílios, ampliando o conjunto de informações já divulgadas sobre os temas no *volume Censo demográfico 2000: primeiros resultados da amostra*.

No que concerne às famílias, as informações divulgadas referem-se ao número de componentes e às características das pessoas residentes em domicílios particulares, como sexo, idade, condição na família, anos de estudo, condição de atividade e de ocupação e rendimento familiar, destacando-se, ainda, a configuração das uniões conjugais a partir da cor das pessoas responsáveis pela família e a cor de seus respectivos cônjuges. As informações sobre os domicílios contemplam aspectos como a existência de iluminação pública, calçamento/pavimentação do logradouro, número médio de moradores por domicílio, condição de ocupação do domicílio, existência de bens duráveis e infra-estrutura de saneamento básico, entre outros indicadores relacionados à adequação das moradias. A publicação inclui, ainda, notas metodológicas e os conceitos e definições utilizados.

O CD-ROM que a acompanha contém, além das informações do volume impresso, dados para as unidades da federação, mesorregiões, microrregiões e municípios, e a íntegra do plano tabular objeto de divulgação do volume *Censo Demográfico 2000: primeiros resultados da amostra*. Os resultados

constituem estimativas obtidas pela expansão da amostra de 5 304 711 domicílios e 20 274 412 pessoas, selecionados em todo o território nacional.

Atlas do censo demográfico 2000

Permite uma interpretação da estrutura e da dinâmica da população brasileira através do mapeamento dos dados coletados no Censo Demográfico 2000, bem como levanta grandes questões que suscitam o debate sobre a população brasileira.

Considerando as diferenças espaciais, apresenta as informações agrupadas nos seguintes temas: o Brasil no contexto mundial; estrutura territorial e demográfica; mobilidade espacial da população e urbanização; condição de habitação e perfil social e econômico da população.

Tabelas, gráficos, fotos, textos e imagens de sensoriamento remoto enriquecem os temas representados cartograficamente. Os dados são apresentados agrupados por grandes regiões, unidades da federação, municípios e ainda em escala intra-urbana.

Metodologia do censo demográfico 2000

Os resultados do Censo Demográfico 2000 revelam o retrato do país no fim do milênio e podem ser considerados a principal fonte de informações para a análise sobre a realidade nacional, servindo para fundamentar o exercício da cidadania, as políticas públicas e os investimentos privados, entre outros aspectos.

Esta publicação apresenta uma descrição detalhada de todos os processos que envolveram a realização desse levantamento, desde as etapas de planejamento e organização da operação, passando pelas novas tecnologias e sistemas desenvolvidos especificamente para a pesquisa, até as diversas formas de divulgação e disseminação dos resultados. Essas informações também estão disponíveis em meio eletrônico, contendo, ainda, a legislação relacionada ao tema, manuais técnicos e gerenciais, instrumentos de coleta, classificações utilizadas, entre outros anexos que complementam e detalham as descrições metodológicas e operacionais.

O IBGE encerra, assim, o plano editorial de divulgação do Censo Demográfico 2000, oferecendo à sociedade não somente um conjunto de resultados através de volumes impressos e arquivos em meio magnético mas também uma completa visão dos procedimentos técnicos e operacionais que levaram à construção de sua base de dados.

13.2.2 Internet

É na Internet que a disseminação de informações encontra sua mídia com maior possibilidade de recursos, visto que com apenas alguns “cliques” pode-se acessar desde uma informação numa tabela estática a complexas tabulações em bancos multidimensionais. Nesse sentido, foi oferecido um amplo conjunto de sistemas e produtos com resultados do Censo 2000 através do portal do IBGE.

Sistema IBGE de Recuperação Automática – SIDRA – esse banco de dados agregados possibilita a montagem de tabelas com os resultados do Censo, bem como a obtenção de *rankings* e a geração de gráficos e cartogramas.

Banco Multidimensional de Estatísticas – BME – pelo endereço www.bme.ibge.gov.br os assinantes podem gerar dados agregados, selecionando variáveis do Censo para execução de cruzamentos a partir dos microdados.

Cidades@ – os resultados do Censo 2000 atualizaram os dados referentes a todos os municípios brasileiros disponibilizados pelo canal Cidades@.

IBGEteen e IBGE 7 a 12 – os canais voltados para adolescentes e crianças, respectivamente, ganharam *links* específicos para divulgar os resultados do Censo 2000 em linguagem adaptada para seu público.

Canal Censo 2000 – além do importante papel que teve na fase de mobilização para o Censo 2000, a página disponibiliza o plano de divulgação, onde é possível acompanhar os lançamentos e o cronograma de novas divulgações.

13.2.3 Arquivos digitais

Em relação aos censos anteriores, a disseminação dos resultados obteve um ganho sobretudo pelas tecnologias de produção e de consulta a informações que surgiram. Os resultados do Censo 2000 puderam ser apresentados em forma de bases de informações georreferenciadas, arquivos de microdados e de dados agregados, malhas digitais de setores censitários e mapas para fins estatísticos. Os produtos digitais foram divulgados em CD-ROM, com geração de cópias sob demanda.

As bases de informações georreferenciadas, ao reunirem resultados do Censo, possibilitaram com o sistema desenvolvido pelo IBGE – *Estatcart: sistema de recuperação de informações georreferenciadas* – a recuperação das informações através de múltiplas opções de consulta e ferramentas para selecionar, localizar e comparar áreas geográficas e variáveis, criar tabelas, indicadores e, ainda, mapas temáticos. As bases por setor censitário possibilitaram, a mais, visualizar mapas de setores censitários urbanos de distrito-sede e sua descrição.

Para preservar a identificação do informante, foram divulgados microdados somente para os resultados da amostra do Censo. Os arquivos foram disponibilizados em formato ASCII (universal).

Para uso em geoprocessamento foram oferecidas malhas de setores censitários rurais e urbanos de acordo com a situação de 1997 da divisão territorial do Brasil adotada pelo Censo. Os mapas para fins estatísticos, municipais e de setores urbanos, e as imagens das folhas de coleta com desidentificação dos informantes, todos em formato pdf, completaram o elenco de serviços.

A seguir, estão listados os produtos do Censo 2000 em CD-ROM, por ordem cronológica de lançamento.

2001

Sinopse preliminar do censo demográfico 2000: sistema de recuperação de informações

Reúne um total de 110 variáveis, abrangendo diferentes recortes geográficos. Permite múltiplas opções de consulta e ferramentas para selecionar, focalizar e comparar áreas geográficas e variáveis, criar tabelas, indicadores através do cruzamento das variáveis apresentadas e, ainda, mapas temáticos do país como um todo ou qualquer outra desagregação geográfica. O usuário pode, ainda, comparar mapas temáticos, imprimir seus resultados ou exportar os dados para planilhas Excel.

Mapas de localidades urbanas para fins estatísticos

São cartogramas e folhas de localidades urbanas para fins estatísticos e mapas de setores censitários urbanos, com limites do perímetro urbano municipal, distrital, subdistrital ou similar e bairro ou similar, sistema viário, identificação das principais edificações e hidrografia e respectivos topônimos, relativos ao distrito-sede de município, vila, área urbana isolada. A escala varia e utiliza a projeção Universal Transversa de Mercator – UTM.

Disponível em papel e em CD-ROM, no formato pdf. Os cartogramas e folhas de localidades urbanas para fins estatísticos não estão disponíveis para o Estado de São Paulo.

Sinopse preliminar do censo demográfico 2000: agregados por setores censitários

Contém informações geradas a partir dos microdados do Censo Demográfico 2000 sobre a população residente por sexo e sobre os domicílios particulares ocupados e não ocupados, em área urbana e rural, como seu tipo e condição de ocupação, para todos os setores censitários do Brasil.

Esses resultados tiveram origem na caderneta do setor censitário, sujeitos a alterações quando da apuração dos questionários preenchidos nos domicílios. Os dados pesquisados foram agregados ao nível de setor censitário da área urbana dos distritos-sede, o nível mais específico, o que corresponde à menor unidade de informação.

As informações são divulgadas para os municípios – de acordo com a divisão territorial do Brasil em 1997 – em dois CD-ROMs: um para municípios com menos de 25 000 habitantes e outro para aqueles com mais de 25 000 habitantes. Inclui os arquivos de descrição dos setores censitários e os códigos e nomes da unidade da federação, mesorregião, microrregião, região metropolitana, município, distrito, subdistrito e bairro.

Base de informações por setor censitário: sinopse preliminar do censo demográfico 2000

Reúne dados preliminares sobre a população e os domicílios do distrito-sede dos municípios. A Base de Informações por Setor Censitário cobre 1 058 municípios com mais de 25 000 habitantes.

Esses resultados tiveram origem na caderneta do setor censitário, sujeitos a alterações quando da apuração dos questionários preenchidos nos domicílios. Os dados pesquisados foram agregados ao nível de setor censitário da área urbana dos distritos-sede, o nível mais específico, o que corresponde à menor unidade de informação.

Disponível em CD-ROM, com visualizador, malhas dos setores censitários digitais e mapas de setores censitários do distrito-sede dos municípios. Requer a instalação do *Estatcart: Sistema de Recuperação de Informações Georreferenciadas*.

2002

Mapas de setores censitários rurais da região metropolitana de São Paulo

São mapas de setores censitários rurais, com limites do setor censitário e sistema viário. Disponíveis em papel e CD-ROM, no formato pdf, em Projeção Universal Transversa de Mercator – UTM –, em escalas distintas, para municípios da região metropolitana de São Paulo.

Mapas municipais estatísticos

São mapas elaborados para cada município brasileiro, apresentando elementos planimétricos, altimétricos e de hidrografia, manchas urbanas, bem como limites internacionais, interestaduais, intermunicipais e limites de áreas especiais, dos distritos e de setores censitários rurais dos municípios.

Disponível em papel e em CD-ROM, no formato pdf, para a maioria dos municípios. Utiliza a projeção Universal Transversa de Mercator – UTM, em escalas que variam de 1:50 000 a 1:250 000. Os mapas apresentam nível de qualidade diferenciado, que deve ser avaliado antes de sua aquisição.

Imagem da folha de coleta do censo demográfico 2000

São arquivos no formato tif, disponibilizados em CD-ROM, com informações cadastrais dos domicílios e unidades não-residenciais de um setor censitário, reunindo: a caracterização dos logradouros do setor, os endereços das unidades visitadas, a espécie da unidade visitada, a numeração das unidades visitadas, a numeração dos domicílios particulares, a numeração dos domicílios coletivos e a indicação dos domicílios particulares selecionados para a amostra.

Base de informações por setor censitário: censo demográfico 2000: resultados do universo

Apresenta os dados relativos aos resultados do universo do Censo Demográfico 2000 através de 520 variáveis selecionadas sobre os domicílios, como abastecimento de água e esgotamento sanitário; sobre a população e os responsáveis pelos domicílios, como sexo e idade; e dados de alfabetização, reunidos no tema população e condições de vida, para os distritos-sede dos municípios que de acordo com a divisão territorial do Brasil de 1997, tinham mais de 25 000 habitantes, o que corresponde a 1 058 municípios.

No caso dos 31 municípios com mais de 500 mil habitantes, os CD-ROMs estão disponíveis para pronta entrega. Para os demais, o prazo de entrega é de quatro dias (96 horas).

Os dados pesquisados através do questionário simplificado do Censo Demográfico 2000 foram agregados ao nível de setor censitário da área urbana dos distritos-sede, o nível mais específico, o que corresponde à menor unidade de informação.

A Base de Informações por Setor Censitário requer a instalação do *Estatcart: sistema de recuperação de informações georreferenciadas*.

Censo demográfico 2000: agregados por setores censitários dos resultados do universo

Contém informações geradas a partir dos microdados do Censo Demográfico 2000, resultados do universo, agregadas ao nível de setor censitário. São 520 variáveis sobre as características das pessoas e dos responsáveis pelos domicílios, como sexo, idade e educação, bem como as características dos domicílios, como o seu tipo, condição de ocupação, forma de abastecimento de água, entre outras.

Para este produto, os dados são divulgados em cinco volumes, sendo um para cada grande região. Inclui notas técnicas e apresenta os códigos e nomes da unidade da federação, mesorregião, microrregião, região metropolitana, município, distrito, subdistrito e bairro.

Malha de setor censitário urbano digital do distrito-sede dos municípios do Brasil: situação 2000

Apresenta a situação em 2000 da divisão político-administrativa do Brasil, através da representação das linhas dos limites que definem os polígonos da maioria dos sub-distritos, bairros e setores censitários urbanos do distrito-sede de 1 058 municípios que tinham mais de 25 mil habitantes em 1996, de acordo com a divisão territorial do Brasil de 1997. A malha está disponível nos formatos AGF e SHAPE, em CD-ROM, em escala variando de 1:5 000 a 1:10 000, com projeção geográfica e Universal Transversa de Mercator - UTM.

Malha de setor censitário rural digital do Brasil: situação 2000

Apresenta a situação em 2000 da divisão político-administrativa do Brasil, através da representação das linhas dos limites que definem os polígonos das grandes regiões, unidades da federação, mesorregiões, microrregiões, municípios, distritos, sub-distritos, setores censitários rurais e áreas urbanas. A malha está disponível nos formatos DGN, DXF e SHAPE, em CD-ROM, nas escalas 1: 500 000, 1:1 000 000 e 1:2 500 000, com projeção geográfica e policônica.

Censo demográfico 2000: microdados da amostra

Apresenta dados sobre a população residente por sexo, cor ou raça, religião, categorias de incapacidade ou deficiência física ou mental, anos de estudo, lugar de nascimento, deslocamentos realizados para o trabalho ou estudo, estado civil, filhos nascidos vivos, óbitos fetais e filhos sobreviventes das mulheres de 10 anos ou mais de idade, grupos ocupacionais, condição de ocupação, atividade econômica e rendimento de trabalho, bem como informações sobre o número de componentes das famílias residentes em domicílios particulares permanentes, existência de bens duráveis e infra-estrutura de saneamento básico nesses domicílios, entre outros aspectos.

Microdados consistem no menor nível de desagregação dos dados de uma pesquisa, retratando, na forma de códigos numéricos, o conteúdo dos questionários, preservado o sigilo das informações. Os microdados possibilitam aos usuários, com conhecimento de linguagens de programação ou *softwares* de cálculo, criar suas próprias tabelas de dados numéricos.

Acompanha os arquivos de microdados uma documentação que fornece nomes e respectivos códigos das variáveis e suas categorias, adicionada dos elementos para o cálculo dos erros amostrais.

2003

Base de informações do censo demográfico 2000: resultados da amostra por município

Apresenta os resultados da amostra do Censo Demográfico 2000 sobre a população, como sexo, cor ou raça, religião, tipo de deficiência física ou mental, anos de estudo, idade, migração e deslocamento, estado civil, atividade econômica e rendimento de trabalho; e sobre os domicílios, como tamanho das famílias, existência de bens duráveis e infra-estrutura de saneamento básico, para Brasil, unidade da federação e município. Este CD-ROM apresenta 197 variáveis relativas aos temas citados. Requer a instalação do *Estatcart: sistema de recuperação de informações georreferenciadas.*

Base de informações do censo demográfico 2000: resultados do universo por município

Apresenta os resultados do universo do Censo Demográfico 2000 sobre a população, relativos a sexo, idade, anos de estudo, estado civil e relação com a pessoa responsável pelo domicílio; e sobre os domicílios, referentes à condição de ocupação, saneamento básico e características das pessoas residentes nos domicílios, em áreas urbana e rural. As informações são divulgadas para Brasil, unidade da federação e município. Este CD-ROM apresenta 512 variáveis relativas aos temas citados. Requer a instalação do *Estatcart: sistema de recuperação de informações georreferenciadas.*

Censo demográfico 2000: agregados por setores censitários dos resultados do universo. 2 .ed.

Reúne dados agregados por setor censitário sobre as características da população, das pessoas responsáveis pelos domicílios e dos domicílios e seus respectivos moradores. A segunda edição do arquivo agregados por setores censitários dos resultados do universo do Censo Demográfico 2000 foi gerada a partir dos microdados do Universo e é composta de 21 planilhas por unidades da federação.

Os dados do cadastro de áreas correspondem aos códigos e nomes dos itens geográficos, além dos totais, médias e variâncias das principais variáveis de dimensão e estratificação dos setores constam de uma planilha MS-EXCEL, por unidade da federação. As informações em nível de setor representam um total de mais de 3 200 variáveis, que abrangem as características dos domicílios, em especial dos domicílios particulares permanentes, bem como dos responsáveis e das pessoas residentes, incorporando os cruzamentos mais solicitados ao IBGE.

Censo demográfico 2000: microdados da amostra. 2. ed.

Apresenta dados sobre a população residente por sexo, cor ou raça, religião, categorias de incapacidade ou deficiência física ou mental, anos de estudo, lugar de nascimento, deslocamentos realizados para o trabalho ou estudo, estado civil, filhos nascidos vivos, óbitos fetais e filhos sobreviventes das mulheres de 10 anos ou mais de idade, grupos ocupacionais, condição de ocupação, atividade econômica e rendimento de trabalho, bem como informações sobre o número de componentes das famílias residentes em domicílios particulares permanentes, existência de bens duráveis e infra-estrutura de saneamento básico nesses domicílios, entre outros aspectos.

A segunda edição apresenta, a mais, informações sobre filhos nascidos vivos por sexo, filhos que estavam vivos por sexo, filhos nascidos mortos por sexo, e sexo do último filho nascido vivo, bem como sobre a existência de identificação, iluminação pública e calçamento/pavimentação no trecho do logradouro onde se localizava o domicílio particular.

Inclui uma nova codificação para o curso mais elevado concluído, conforme a Comissão Nacional de Classificação – CONCLA e, ainda, a codificação da ocupação e da atividade utilizadas no Censo Demográfico 1991.

Também foi modificada a codificação para os países, conforme a estrutura estabelecida pela Organização das Nações Unidas – ONU.

Censo demográfico 2000: microdados da amostra: amostra de uso público

Apresenta dados para uma subamostra dos registros da amostra do Censo Demográfico com um tamanho adequado para processamento em microcomputadores comumente utilizados nessas tarefas.

O CD-ROM com os registros da amostra de uso público contém, ainda, um programa instalador que ao ser executado gravará no microcomputador do usuário arquivos em formato txt – contendo os registros com as variáveis das unidades domiciliares da amostra de uso público e registros com as variáveis dos respectivos moradores – e arquivos em formato excel, contendo planilhas com informações sobre o número de domicílios e de pessoas na amostra de uso público e no universo para cada estrato definido no plano amostral.

Inclui arquivos com descrição da amostra de uso público, definição das variáveis e as classificações utilizadas.

Base de informações por área de ponderação: censo demográfico 2000: resultados da amostra.

Reúne dados do Censo Demográfico 2000, obtidos através do questionário da amostra, para os 484 municípios que tiveram mais de uma área de ponderação. São 182 variáveis selecionadas, referentes a características gerais da população, educação, migração, deslocamento, nupcialidade, fecundidade, mortalidade infantil, trabalho, rendimento, famílias e domicílios.

Define-se Área de Ponderação como sendo a menor unidade geográfica para divulgação dos resultados da amostra do Censo Demográfico 2000, formada por um agrupamento de setores censitários.

A Base de Informações por Área de Ponderação requer a instalação do *Estatcart: sistema de recuperação de informações georreferenciadas*.

Cartograma municipal dos setores censitários: situação 2000

Retrata a situação vigente da divisão político-administrativa do País, através da representação vetorial dos limites dos municípios, distritos, sub-distritos e setores censitários rurais, na escala em que foram elaborados os mapas municipais estatísticos, podendo variar de 1:50 000 a 1:250 000, de acordo com o mapeamento topográfico existente na época da elaboração dos mapas municipais. As bases cartográficas, em formato vetorial, são constituídas por linhas, cujos pontos foram definidos em coordenadas geográficas - latitude e longitude - expressas em unidades de grau sexagesimal, e decimais de grau. Os arquivos estão no formato DGN.

Referências

CATÁLOGO DO IBGE 2003. Rio de Janeiro: IBGE, 2003. 214 p.

IBGE. Rio de Janeiro, 2003. Disponível em: <http://www.ibge.gov.br>. Acesso em: 9 out. 2003.

RELATÓRIO de avaliação da etapa de coleta do Censo 2000. Rio de Janeiro, 2001. 22 p.

14 Base de dados

Para o armazenamento dos dados do Censo Demográfico 2000, durante o processo de produção, optou-se pelo Statistical Analysis System - SAS em ambiente *mainframe* IBM OS/390, após estudo comparativo com o sistema ORACLE.

Para a disseminação dos microdados, foi desenvolvido um banco de dados multidimensional, denominado Banco Multidimensional de Estatísticas – BME, utilizando-se SGBD ORACLE em ambiente IBM/AIX.

Para a disseminação dos dados agregados, optou-se pelo uso de um sistema já existente, o Sistema IBGE de Recuperação Automática – SIDRA , utilizando-se o SGBD ORACLE em ambiente Windows 2000.

Tanto o BME quanto o SIDRA estão disponíveis no portal do IBGE, na Internet.

14.1 Banco Multidimensional de Estatísticas

Visando facilitar o acesso, a recuperação e o manuseio dos dados estatísticos, foi criado o Banco Multidimensional de Estatísticas - BME, a partir das pesquisas produzidas pelo IBGE. O banco de dados BME é formado por microdados (dados coletados nos questionários das pesquisas). O resultado de uma consulta no BME é uma tabela agregada onde é preservado o sigilo legal da informação estatística.

14.1.1 Objetivo do projeto

O Projeto BME tem como objetivo tornar disponível ao público ferramentas voltadas à busca, recuperação e manuseio das informações estatísticas sobre uma base de microdados – base de dados totalmente desagregada. Desta forma, o BME introduz uma revisão conceitual na organização da base de dados do IBGE, com a incorporação de métodos seguros para o manuseio de informações desagregadas, por meio de um serviço ativo de metainformação.

14.1.2 Importância da tecnologia e suas facilidades

O Projeto BME incorporou a tecnologia corrente para manuseio de grandes volumes de informações. O conceito de dados multidimensionais foi implementado com o modelo *Relational On-line Analytical Processing* – ROLAP sobre uma base de dados relacional. Um metamodelo é utilizado para armazenar as informações sintáticas e semânticas necessárias à busca, recuperação e manuseio das informações. As ferramentas de acesso aos dados utilizam este metamodelo para estruturar as consultas e fazer um manuseio inteligente da informação.

A interface Web do BME permite utilizar os mecanismos de busca sobre a metainformação, visando localizar variáveis, exibir conceitos ou conhecer planos de classificação das informações. A navegação necessária para recuperar e agregar as informações é efetuada *on-line,* de forma transparente, sem necessidade de codificar programas. Os usuários contam com um mecanismo para elaboração de consultas, que ajuda a prevenir falhas e a manusear informações heterogêneas. A partir deste mecanismo, as informações esparsas são consideradas apenas nos anos em que ocorreram, e as informações distribuídas, por diversas tabelas, são reunidas, automaticamente, de acordo com regras previamente definidas. As informações amostrais são, automaticamente, expandidas.

O bom desempenho do processo de recuperação de valores permite aos usuários a realização de análises interativas sobre os dados, possibilitando um refinamento gradual das informações obtidas.

Com o BME, é possível manusear variáveis quantitativas, calcular freqüências de cruzamentos de classificações e definir novas variáveis, derivadas das variáveis armazenadas, para ampliar uma investigação.

14.1.3 Resultados obtidos

A base de dados do BME é formada por microdados, originados no Censo Demográfico 2000 e pela metainformação associada a sua produção.

Os dados armazenados, chamados microdados, correspondem aos dados existentes no Questionário Básico – CD 1.01, Questionário da Amostra – CD 1.02 e na Folha de Coleta – CD 1.07. Visando facilitar o manuseio dos microdados são, também, acrescentadas algumas outras variáveis derivadas.

Todos as variáveis pesquisadas no Censo Demográfico 2000 estão disponibilizadas no BME, agrupadas por temas e sub-temas. Visando facilitar o acesso e o tratamento dos dados, alguns quesitos são replicados, existindo, simultaneamente, em temas diversos. Além disso, para cada variável quantitativa, foi criada uma variável categorizada com uma ou mais classe de categorias, por exemplo:

Idade em anos

Idade da pessoa em anos completos na data de referência da pesquisa.

Idade em anos, classe

Idade da pessoa em anos completos na data de referência da pesquisa, em classes de valores.

Classificação da informação:

1 - De 0 a 4 anos

2 - De 5 a 9 anos

3 - De 10 a 14 anos

4 - De 15 a 19 anos

5 - De 20 a 24 anos

6 - De 25 a 29 anos

7 - De 30 a 34 anos

8 - De 35 a 39 anos

9 - De 40 a 44 anos

10 - De 45 a 49 anos

11 - De 50 a 54 anos

12 - De 55 a 59 anos

13 - De 60 a 64 anos

14 - De 65 a 69 anos

15 - De 70 a 74 anos

16 - De 75 a 79 anos

17 - De 80 a 84 anos

18 - De 85 a 89 anos

19 - De 90 a 94 anos

20 - De 95 a 99 anos

21 - 100 anos ou mais

Estrutura de temas e sub-temas:

Dados da Amostra

- **Domicílios**
  - Bens duráveis
  - Características gerais
  - Cômodos
  - Marcas de imputação no registro de domicílios
  - Moradores
  - Rendimento domiciliar
  - Saneamento básico
- **Famílias**
  - Componentes da família
  - Rendimentos da família
- **Pessoas**
  - Características gerais
  - Deficiência física ou mental
  - Fecundidade
    - Último filho nascido vivo
    - Total de filhos
  - Instrução
  - Migração
  - Nupcialidade
  - Previdência
  - Rendimentos
  - Trabalhos
    - Demais trabalhos
    - Todos os trabalhos
    - Trabalho principal

Dados do Universo

- **Domicílios**
  - Características gerais
  - Marcas de imputação no registro de domicílios
  - Moradores
  - Saneamento básico

Pessoas

Marcas de imputação no registro de pessoas

Pessoas responsáveis pelo domicílio e moradores individuais

Rendimento por morador

14.1.4 Segurança dos dados

O ambiente multidimensional contém uma administração centralizada que assegura:

- identificação de usuários através de perfis preestabelecidos;
- capacidade de recuperação e segurança física dos dados;
- controle e registro do acesso aos dados; e
- alto nível de automação das funções de administração dos serviços.

14.1.5 Acesso aos dados

O BME é dirigido ao público em geral e, em especial, aos profissionais envolvidos em tarefas de análise de informações. Os usuários do BME manuseiam diretamente o serviço, através de um *browser* conectado à Internet, onde podem conhecer os conceitos apresentados na metainformação das pesquisas e executar seus próprios cruzamentos de dados.

O acesso aos dados pode ser feito para obter um cruzamento de informações. Esta função é utilizada depois de ser selecionado, pelo menos, um item de informação. Para execução da consulta são oferecidos filtros, temporal e espacial, e uma ferramenta de ordenação das variáveis. Caso não sejam utilizados, o BME executa a consulta com os dados mais recentes, no maior nível espacial de agregação, e cruza os dados na ordem em que as variáveis foram selecionadas. Estão, ainda, disponíveis filtros sobre os dados manuseados.

O BME dispõe de uma segunda modalidade de manuseio dos dados, onde o acesso é feito a resultados de consultas pré-definidas, isto é, tabelas já existentes no plano de divulgação. Nesta segunda modalidade, os dados são oferecidos de forma agregada, permitindo ao usuário maior rapidez no acesso a uma informação ou a uma tabela de informações. Os dados apresentados podem ser agrupados, restringindo-se as dimensões manuseadas. O sistema também permite que uma informação agregada possa ser detalhada na modalidade de acesso a “Consultas Livres”, para onde é transferido o contexto da informação desejada.

14.2 Sistema IBGE de Recuperação Automática

O Sistema IBGE de Recuperação Automática – SIDRA foi pioneiro na disseminação de dados agregados no IBGE. Originário da década de 80 e desenvolvido para computadores de grande porte (*mainframes*), foi sendo aperfeiçoado ao longo dos anos, de modo a se ajustar às novas tecnologias de comunicação e armazenamento de dados, incorporadas à arquitetura de informática do

IBGE (Telex, Rede Nacional de Comunicação por Comutação de Pacotes – RENPAC, sistemas de gerenciamento de arquivos e de banco de dados etc.).

14.2.1 Objetivo do projeto

A idéia do Banco de Dados Agregados do SIDRA é armazenar as tabelas, contendo os dados agregados provenientes dos planos tabulares das pesquisas do IBGE. As tabelas são previamente definidas pelos técnicos responsáveis pelas pesquisas.

Com o advento da Internet, tornou-se imprescindível a modernização do sistema para que pudesse ser acessado por esse meio. Assim, em 1997, o IBGE lançou uma primeira versão do SIDRA para a Internet e, em abril de 2001, lançou a segunda versão, atualmente em uso, que permitiu o armazenamento e a consulta aos dados agregados do Censo 2000 pela Internet.

14.2.2 Importância da tecnologia e suas facilidades

O SIDRA incorpora inovações no uso da Tecnologia de Informação no IBGE. Opera num servidor com ambiente operacional Windows 2000 Server, com Internet Information Services – IIS 5, Office 2000 Web Components, Internet Map Server – Arc IMS, e utiliza o SGBD Oracle 8i. Foi implementado utilizando ferramentas para o desenvolvimento de aplicações Web: Visual Basic, ASP (VB Script), Visual InterDev e Visual SourceSafe.

As principais facilidades de consulta disponíveis são:

Navegação pelo acervo – permite ao usuário a navegação hierárquica pelos diversos tipos de descritores das tabelas (área temática, pesquisa, variável, período, nível territorial etc.), de modo a encontrar a tabela que possui os dados desejados;

Procura por palavra-chave – especificando uma ou mais palavras, é exibida a lista das tabelas que contém as palavras digitadas;

Acesso rápido aos dados da tabela – basta digitar o número da tabela para acessar os seus dados;

Consulta por temas – apresenta um resumo dos dados disponíveis associados ao tema;

Lista conjuntural – possibilita receber por *e-mail* um resumo dos dados das pesquisas conjunturais, uma vez divulgados, assim que estes são carregados no banco de dados;

Páginas de pesquisas – quando do lançamento dos dados de novas pesquisas no banco, são elaboradas páginas específicas que descrevem os novos dados disponíveis;

Montagem de tabela – é possível customizar a tabela. O usuário seleciona os dados que deseja e de que forma, se nas linhas, colunas ou no cabeçalho da tabela resultante;

Ranking – o usuário pode obter um *ranking*, ordenando uma das dimensões da tabela, de modo a saber, por exemplo, quais são os municípios mais populosos de determinada Unidade da Federação;

Gráficos – é possível gerar mais de 20 tipos de gráficos; e

Cartogramas – o usuário pode comandar a geração de cartogramas (mapas temáticos) para representar, espacialmente, uma determinada informação da tabela (por exemplo, a população residente de 60 anos ou mais de idade nos municípios do Rio de Janeiro).

14.2.3 Resultados obtidos

O conjunto de tabelas de uma pesquisa carregado no Banco de Dados Agregados compõe o que é chamado de plano tabular. Para cada fase de liberação de resultados do Censo Demográfico 2000 (Resultados Preliminares, Sinopse Preliminar, Resultados do Universo e Resultados Preliminares da Amostra), foram criados um plano tabular e uma página na área de "Páginas de Pesquisas" do *site* do SIDRA, comemorativa do lançamento dessas informações no banco de dados. Os Resultados Preliminares foram disseminados no formato de planilhas Excel prontas.

A disseminação de tabelas de dados agregados através do SIDRA apresenta vantagens, comparada às mídias convencionais (papel e CD-ROM): inexistência de um formato pré-definido (o usuário escolhe como quer a tabela); e a falta da limitação física do volume de papel ou CD-ROM, o que permite uma grande desagregação espacial (dados por município, distrito e bairro, por exemplo) para a quase totalidade das tabelas.

Outro resultado importante do Banco de Dados Agregados é a possibilidade de criação de séries temporais, o que permite a comparabilidade dos dados ao longo do tempo. Para o Censo Demográfico 2000, procurou-se, sempre que possível, organizar os dados de modo a favorecer a comparação com os obtidos no Censo Demográfico 1991.

14.2.4 Segurança dos dados

Os dados estão armazenados em banco ORACLE e submetidos às rotinas de segurança da área de serviços da Diretoria de Informática, o que garante sua recuperação e segurança física.

O acesso à aplicação de consulta é controlado mediante recursos de um *firewall,* de modo a bloquear o acesso via Internet a serviços de uso interno da instituição.

14.2.5 Acesso aos dados

O SIDRA é dirigido ao público em geral e, em particular, aos técnicos do IBGE envolvidos em tarefas de análise de informações das pesquisas.

O manuseio do SIDRA é feito diretamente pelos usuários através de um *browser* conectado à Internet, onde podem conhecer os conceitos apresentados na metainformação das pesquisas e recuperar as tabelas incorporadas ao Banco de Dados Agregados.

14.3 Escolha do SAS

Conforme diretriz institucional, o SAS é a principal ferramenta de análise estatística utilizada pelos técnicos do IBGE. Partindo desta premissa, concentraram-se esforços para maximizar o uso do ambiente SAS, particularmente em seus recursos para análise de dados. Assim, com relação aos dados do Censo Demográfico 2000, criou-se uma base de dados no *mainframe*, utilizando-se o SAS.

Os arquivos SAS foram criados com as características de um banco de dados, de acordo com regras de integridade e de construção de índices.

Com relação ao acesso aos dados, foram desenvolvidas interfaces gráficas para ambiente de microcomputadores, e disponibilizaram-se os dados para serem acessados em ambiente *mainframe,* mediante programas SAS a serem desenvolvidos pelos técnicos do IBGE.

14.3.1 Escolha do *software*

A base de dados do Censo Demográfico 2000 é composta por dados do universo, por dados da amostra e por variáveis derivadas. A base foi concebida de forma a atender aos seguintes objetivos: recuperação rápida das informações e fácil acesso aos dados pelos técnicos do IBGE.

Para atender aos objetivos estipulados, foi realizado um estudo comparativo, com base nos dados do Censo Demográfico 1991, entre as formas de armazenamento do SAS e do ORACLE. Para a escolha do banco de dados, foram realizados testes comparativos entre o armazenamento em SAS e ORACLE, utilizando-se dados do universo (8 268 124 domicílios e 31 588 925 pessoas) do Estado de São Paulo, referentes ao Censo Demográfico 1991, com o objetivo de verificar os seguintes pontos:

- tempo de acesso aos dados;
- segurança dos dados;
- facilidades no manuseio dos arquivos (inclusão, alteração e exclusão de dados); e
- interfaces de acesso, tanto em microcomputador como no *mainframe.*

Após comparar os resultados dos testes, foi possível concluir que os resultados obtidos, utilizando a base SAS como banco de dados, geraram os melhores tempos de acesso e processamento. Portanto, foi a forma de armazenamento escolhida para o Censo Demográfico 2000.

14.3.2 Montagem das bases de dados

Os arquivos de dados do Censo Demográfico 2000 foram divididos em dois tipos, a saber: dados do universo e dados da amostra. Para cada tipo de arquivo, foram feitas duas subdivisões: dados não criticados e dados criticados. A seguir, é feita uma breve descrição de cada um deles.

Arquivos com dados do universo não criticados

Foram gerados arquivos de dados do universo para cada Unidade da Federação, perfazendo um total de 27 arquivos, contendo as informações de controle, de Domicílio e de Pessoa. Estes dados vieram direto da captura de dados.

Arquivos com dados do universo criticados

Estes arquivos passaram por processos de crítica e imputação, gerando no total 27 arquivos, um para cada Unidade da Federação. Nestes arquivos foram adicionadas variáveis derivadas.

Arquivos com dados da amostra não criticados

Os arquivos com os dados da amostra foram gerados para cada Unidade da Federação, totalizando 27 arquivos, contendo as informações de controle, os quesitos de Domicílio e os quesitos de Pessoa, e outras variáveis adicionais da amostra.

Arquivos com dados da amostra criticados

Os arquivos dos dados da amostra, vindos da coleta, passaram por processos de crítica e imputação, gerando 27 arquivos, um para cada Unidade da Federação. Foram adicionadas, nestes arquivos, variáveis derivadas e as referentes ao procedimento de expansão da amostra.

14.3.3 Segurança dos dados

Foi utilizado o produto *Resource Access Control Faciliy* – RACF, do *mainframe* IBM 9672, para que somente técnicos do IBGE autorizados tivessem acesso aos arquivos SAS do Censo Demográfico 2000.

O RACF é um produto IBM que implementa um mecanismo de controle para ajudar a gerenciar o acesso à informação. Além de garantir a integridade física dos arquivos (proteção contra gravação e exclusão por siglas não autorizadas), disponibiliza relatórios que mostram as operações realizadas pelos técnicos do IBGE.

14.3.4 Acesso aos dados

Para o acesso interno dos técnicos do IBGE aos dados do Censo Demográfico 2000, foram disponibilizadas três alternativas, descritas abaixo:

a) programação SAS (convencional), em ambiente Windows ou *mainframe* – os técnicos do IBGE, de posse da documentação gerada pela base de dados, obtêm a descrição dos registros (física e lógica) e os nomes dos arquivos a serem acessados;

b) banco de Metadados – na estação de trabalho, conforme sua necessidade, o usuário comanda a geração de um *book SAS*, que o habilita a acessar os dados, através de um programa SAS. Cabe ressaltar que, nesta funcionalidade, não estão descritos os conceitos das variáveis, estando apenas disponíveis informações do tipo: nome da variável, formato, tamanho, posição etc.; e

c) interface cliente/servidor – executada dentro do ambiente SAS, que gera um arquivo SAS com os dados resultantes da solicitação do usuário. Essa interface foi construída na linguagem SCL do SAS.

Equipe técnica

Grupo de trabalho

Antonio Carlos Fernandes de Menezes – consultor
Carmen Zagari Machado
Cássia Rezende Pinho
Lana Lima Moreira
Maria Vilma Salles Garcia
Paulo Cesar de Souza Quintslr
Sonia Albieri
Sonia Regina Allevato
Valéria Grace Costa

Revisão gramatical

Ceni Maria de Paula de Souza
Elizabeth Saramela Nogueira
Iaracy Prazeres Gomes
José Antônio de Almeida Senna
Letícia Maria Acioly de Lima
Paulo Roberto Fiore de Castro
Robson Waldhelm
Rosane Tavares Trindade
Rosemere Nunes Ventura

Consolidação e revisão final

Antonio Carlos Fernandes de Menezes – consultor
Sonia Albieri

Presidência

Redação

Luiz Mario Gazzaneo
Maria Vilma Salles Garcia
Vera Regina Affonso de Oliveira

Revisão

Maria Vilma Salles Garcia

Diretoria de Pesquisas

Redação

Afonso Celso Calvo Rangel
Alexandre dos Reis Santos
Alicia Bercovich
Ana Lúcia Saboia
Ana Paula Moura Reis Miceli
Andréa da Silva Borges
Andréa Diniz da Silva
Andréa Machado Barbosa
Angela Filgueiras Jorge
Antonio Carlos Fernandes de Menezes – consultor
Antonio José Ribeiro Dias
Ari do Nascimento Silva
Bruno Freitas Cortez
Celso Cardoso da Silva Simões
Djalma Galvão Carneiro Pessoa
Eliane Aparecida de Araújo Xavier
Elisa Lustosa Caillaux
Eneiza de Andrade Ferreira
Ennio Leite de Mello
Fernando Roberto Pires de Albuquerque
Geraldo José Polidoro
Guilherme Guimarães Moreira
Jacqueline dos Santos Manhães Pinto
Juarez de Castro Oliveira
Laura Baridó Indá
Leila Regina Ervatti
Luís Carlos de Souza Oliveira
Magdalena Cronemberger Góes
Márcia Regina Martins Lima Dias
Marco Antonio dos Santos Alexandre
Maria Dolores Bombardelli Kappel
Neimar Rodrigues Guimarães
Nilza de Oliveira Martins Pereira
Rodrigo Aires Lemes
Rosemary Vallejo de Azevedo
Sonia Albieri
Tereza Cristina Nascimento Araújo
Vandeli dos Santos Guerra

Revisão

Alicia Bercovich
Antonio Carlos Fernandes de Menezes – consultor
Antonio José Ribeiro Dias
Ari do Nascimento Silva
Eliane Aparecida de Araújo Xavier
Francisco de Assis Correa Alchorne
Joana D'Arc Corrêa Marques
Jorge Mendes Carneiro
Laura Baridó Indá
Luiz Alberto Matzenbacher
Marco Antonio dos Santos Alexandre
Sonia Albieri
Sonia Regina Madeira
Vera Lúcia Manfredini Rodrigues de Lima
Vilma de Oliveira Campos Peixoto
Zélia Magalhães Bianchini

Apoio técnico

Alessandra Soares da Poça
Antonio Roberto Pereira Garcez
Carlos Renato Caldeira Grilo
Cézar Cioffi Camardella
Cláudia Bahia de Araujo
Debora Ferreira de Souza
Eudes dos Santos Monteiro Junior
Helem Ortega da Silva
Janaína Reis Xavier Senna
Jorge da Silva
José Ademir Campos de Carvalho
José Luís Petruccelli
Kátia de Fátima Dias
Letícia Maria Brandão e Souza
Luciana Martins Gomes
Luciano Tavares Duarte
Márcia Luzia Coenca Maia
Marcos Ribeiro de Mattos
Maria de Fátima Cortezia Coelho
Maria de Fatima Lobo Augusto
Mário Fernandes Filho
Mauro Sorge
Nadia Regina Paiva de Souza
Priscila Carvalho de Araujo Mesquita
Renato D'Almeida Cunha Bastos
Ricardo Luiz Cardoso
Romeu Ferreira Emygdio
Sandra Cavalcanti de Barros
Therezinha Maria Lamêgo do Nascimento
Valéria Martins Bourguignon Beiriz
Viviane Cirillo Carvalho Quintaes

Diretoria de Geociências

Redação

Angela Maria de Souza Ferreira
Carmen Zagari Machado
Cristiane Moreira Rodrigues
Dulce Santoro Mendes
Paulo Cesar Martins
Sonia Luiza Terron
Valeria Grace Costa
Wolmar Gonçalves Magalhães
Wolney Cogoy de Meneses

Revisão

Carmen Zagari Machado
Moema José de Carvalho Augusto
Sérgio Medeiros de Lavour
Valeria Grace Costa
Wolney Cogoy de Meneses

Diretoria de Informática

Redação

Antonio José de Oliveira
Ataíde José de Oliveira Venâncio
Cláudio Mariano Fernandes
Heleno Ferreira Mansoldo
Horacio Coelho da Costa Ferreira
Luiz Antonio Gauziski de Araujo Figueiredo

Revisão
Cássia Rezende de Pinho
Edgar Anzanello
Heleno Ferreira Mansoldo
Maria Célia Pelisson Jacon
Reina Marta Hanono

Diretoria Executiva

Redação
Cidália Maria Conceição Moreira Augusto
Franklin Moreira de Almeida
Josiane Heil Figueira
Lana Lima Moreira

Revisão
Lana Lima Moreira
Nuno Duarte Bittencourt

Centro de Documentação e Disseminação de Informações

Redação
Carlos José Lessa de Vasconcelos
David Wu Tai
Edna Campello
Izabelle de Oliveira
Leda Pereira da Silva
Márcio Imamura
Mário Henrique Monteiro Mattos
Marise Maria Ferreira
Paulo César de Souza Quintslr
Rose Maria Barros de Almeida da Silva
Solange Makrakis

Revisão
Paulo César de Souza Quintslr
Rose Maria Barros de Almeida da Silva

Projeto editorial

Coordenação de Produção
Marise Maria Ferreira

Gerência de Editoração

Estruturação textual, tabular e de gráficos
Beth Fontoura
Carmen Pessoa
Katia Vaz Cavalcanti

Copidesque e revisão
Anna Maria dos Santos
Cristina R. C. de Carvalho
Kátia Domingos Vieira
Maria de Lourdes Amorim
Sueli Alves de Amorim

Diagramação tabular
Beth Fontoura
Neuza Damásio
Sônia Gonçalves Rocha

Diagramação textual e de gráficos
Carlos Amaro Feliciano
Lgonzaga
Maria da Graça Fernandes Lima

Programação visual
Luiz Carlos Chagas Teixeira

Gerência de Gráfica

Impressão
José Augusto dos Santos

Gerência de Documentação

Normalização bibliográfica
Ana Raquel Gomes da Silva
Angélica Sodré dos Santos
Aparecida Tereza Rodrigues Regueira
Diva de Assis Moreira
Eliane Maricá Soares
Elisabete Siqueira Soares
Michelle Cristina Romualdo de Souza
Sonia Regina Allevato

Gráfica Digital

Impressão e acabamento
Ednalva Maia do Monte

Série Relatórios Metodológicos

ISSN 0101-2843

Números Divulgados

volume 1 - Metodologia da Pesquisa Nacional por Amostra de Domicílios na década de 70, ISBN 85-240-0005-8,1981

volume 2 - Metodologia da Pesquisa Mensal de Emprego - 1980, ISBN 85-240-0129-1, 1983

volume 3 - Metodologia das Pesquisas Agropecuárias Anuais - 1981: Produção Agrícola Municipal, Produção da Pecuária Municipal, Produção Extrativa Vegetal, Silvicultura, ISBN 85-240-0132-1, 1983

volume 4 - Metodologia do Censo Demográfico de 1980, ISBN 85-240-0131-3, 1983

volume 5 - Metodologia do Censo Agropecuário de 1980, ISBN 85-240-0229-8, 1985

volume 6 - Pesquisas Agropecuárias, ISBN 85-240-0305-7, 1989

volume 7 - Matriz de Insumo-Produto - Brasil - 1980, ISBN 85-240-0307-3, 1989

volume 8 - Sistema de Contas Nacionais Consolidadas - Brasil, ISBN 85-240-0319-7, 1989

volume 9 - Produto Interno Bruto - Brasil, ISBN 85-240-0325-1, 1989

volume 10 - Pesquisa de Orçamentos Familiares, ISBN 85-240-0361-8

v.1 - Obtenção das Informações em Campo, ISBN 85-240-0359-6, 1990

v.2 - Tratamentos das Informações, ISBN 85-240-0358-8, 1991

v.3 - Aspectos de Amostragem, ISBN 85-240-0360-X, 1991

volume 11 - Indicadores Conjunturais da Indústria: produção, emprego e salário, ISBN 85-240-0352-9, 1991

volume 12 - Pesquisa Anual de Comércio - PAC, ISBN 85-240-0403-7, 1991

volume 13 - Pesquisa Anual do Transporte Rodoviário - PATR, ISBN 85-240-0405-3, 1991

volume 14 - Sistema Nacional de Preços ao Consumidor: métodos de cálculo, ISBN 85-240-0495-9, 1994

volume 15 - Pesquisa Mensal de Comércio - PMC, ISBN 85-240-0608-0, 1996

2ª edição, ISBN 85-240-3674-5, 2003

volume 16 - Sistema Nacional de Índices de Preços ao Consumidor: Ajustamento Sazonal, ISBN 85-240-0625-0, 1996

volume 17 - Pesquisa Industrial Anual e Pesquisa Anual da Indústria da Construção - PIA e PAIC, ISBN 85-240-0636-6, 1997

volume 18 - Matriz de Insumo-Produto, ISBN 85-240-0654-4, 1997

volume 19 - Produto Interno Bruto Trimestral, ISBN 85-240-0754-0, 1999

volume 20 - Regionalização das Transações do Setor Público, ISBN 85-240-0757-7, 2000

volume 21 - Sistema Nacional de Índices de Preços ao Consumidor: Estruturas de Ponderação a partir da Pesquisa de Orçamentos Familiares 1995-1996, ISBN 85-240-0766-4

v.1 - Metodologia, ISBN 85-240-0765-6, 2000

v.2 - Estruturas de ponderação, pesos regionais e tradutor, ISBN 85-240-0764-8, 2000

volume 22 - Estimativas da População do Brasil, Grandes Regiões, Unidades da Federação e Municípios, ISBN 85-240-3070-4, 2002

volume 23 - Pesquisa Mensal de Emprego, ISBN 85-240-3081-X, 2002

volume 24 - Sistema de Contas Nacionais, ISBN 85-240-3702-4, 2003

volume 25 - Metodologia do Censo Demográfico 2000, ISBN 85-240-3700-8, 2003